Osvaldo Gervasi · Beniamino Murgante ·
Chiara Garau · Ana Maria A. C. Rocha ·
M. Noelia Faginas Lago · Ana C. Braga ·
Yeliz Karaca (Eds.)

LNCS 16763

Computational Science and Its Applications – ICCSA 2026 Workshops

Braga, Portugal, June 30 – July 3, 2026
Proceedings, Part VI

6 Part VI

Springer

Lecture Notes in Computer Science **16763**

Founding Editors

Gerhard Goos
Juris Hartmanis

Editorial Board Members

Elisa Bertino, *Purdue University, West Lafayette, IN, USA*
Wen Gao, *Peking University, Beijing, China*
Bernhard Steffen, *TU Dortmund University, Dortmund, Germany*
Moti Yung, *Columbia University, New York, NY, USA*

The series Lecture Notes in Computer Science (LNCS), including its subseries Lecture Notes in Artificial Intelligence (LNAI) and Lecture Notes in Bioinformatics (LNBI), has established itself as a medium for the publication of new developments in computer science and information technology research, teaching, and education.

LNCS enjoys close cooperation with the computer science R & D community, the series counts many renowned academics among its volume editors and paper authors, and collaborates with prestigious societies. Its mission is to serve this international community by providing an invaluable service, mainly focused on the publication of conference and workshop proceedings and postproceedings. LNCS commenced publication in 1973.

Osvaldo Gervasi · Beniamino Murgante ·
Chiara Garau · Ana Maria A. C. Rocha ·
M. Noelia Faginas Lago · Ana C. Braga ·
Yeliz Karaca
Editors

Computational Science and Its Applications – ICCSA 2026 Workshops

Braga, Portugal, June 30 – July 3, 2026
Proceedings, Part VI

Springer

Editors
Osvaldo Gervasi
University of Perugia
Perugia, Perugia, Italy

Chiara Garau
University of Cagliari
Cagliari, Italy

M. Noelia Faginas Lago
University of Perugia
Perugia, Italy

Yeliz Karaca
University of Massachusetts
Worcester, MA, USA

Beniamino Murgante
University of Basilicata
Potenza, Italy

Ana Maria A. C. Rocha
University of Minho
Braga, Portugal

Ana C. Braga
University of Minho
Braga, Portugal

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Computer Science
ISBN 978-3-032-30529-9 ISBN 978-3-032-30530-5 (eBook)
https://doi.org/10.1007/978-3-032-30530-5

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Switzerland AG 2027

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

Preface

These 9 volumes (LNCS volumes 16758–16766) comprise the peer-reviewed papers from the 57 Workshops of the 2026 International Conference on Computational Science and Its Applications (ICCSA 2026), held between June 30 and July 3, 2026, in Braga, Portugal. The peer-reviewed papers from the six main conference tracks are published in a separate set of 3 volumes (LNCS volumes 16767–16769).

The conference was held in a hybrid format, with the vast majority of participants attending in person, hosted by the University of Minho in Braga, Portugal.

We enabled virtual participation for those who did not attend the event in person due to logistical, political, or economic constraints by adopting a technological infrastructure based on open-source software (Jitsi + Riot) and a commercial Cloud infrastructure. ICCSA 2026 was organized by the University of Minho (Portugal), the University of Perugia (Italy), the University of Basilicata (Italy), the University of Cagliari (Italy), Monash University (Australia), and Kyushu Sangyo University (Japan).

With the 2026 edition, ICCSA celebrated its 26th anniversary, marking a new chapter in its longstanding history after its memorable 25th anniversary celebrated in Istanbul, Türkiye. Hosted in Braga, Portugal, the conference found a particularly meaningful setting in one of Europe's oldest and most vibrant academic cities, a place where historical depth, cultural continuity, and contemporary scientific exchange meet. Braga offers an inspiring European context for renewing the international and intercultural dialogue that has always characterized ICCSA, while giving new visibility to the role of computational science in addressing the complex challenges of cities, territories, technologies, and societies.

ICCSA 2026 marked another fruitful and thought-provoking academic event in the International Conference on Computational Science and Its Applications (ICCSA) conference series, previously held in Istanbul, Türkiye (2025), Hanoi, Vietnam (2024), Athens, Greece (2023), Málaga, Spain (2022), Cagliari, Italy (hybrid with few participants in presence in 2021 and completely online in 2020), Saint Petersburg, Russia (2019), Melbourne, Australia (2018), Trieste, Italy (2017), Beijing, China (2016), Banff, Canada (2015), Guimarães, Portugal (2014), Ho Chi Minh City, Vietnam (2013), Salvador, Brazil (2012), Santander, Spain (2011), Fukuoka, Japan (2010), Suwon, South Korea (2009), Perugia, Italy (2008), Kuala Lumpur, Malaysia (2007), Glasgow, UK (2006), Singapore (2005), Assisi, Italy (2004), Montreal, Canada (2003), and (as ICCS) Amsterdam, the Netherlands (2002) and San Francisco, USA (2001).

Computational Science constitutes the main pillar of most present research, industrial, and commercial applications, and plays a unique role in exploiting innovative ICT technologies, and the ICCSA conference series has, accordingly, provided ample opportunities to researchers and industry practitioners to discuss new ideas, to share complex problems and their solutions, and to shape new trends in Computational Science. As the conference mirrors society from a scientific point of view, this year's undoubtedly dominant theme was large language models, machine learning, and Artificial Intelligence,

and their applications across the most diverse technological, economic, and industrial fields, amongst others.

The ICCSA 2026 conference included 57 workshops, focusing on topical issues of utmost importance to science, technology, and society: from new computational approaches for earth science, to mathematical methods for image processing, new statistical and optimisation methods, several Artificial Intelligence approaches, sustainability issues, smart cities and related technologies, to name a few. In addition, ICCSA 2026 included six general tracks covering computational science and its applications: Computational Methods, Algorithms and Scientific Applications; High-Performance Computing and Networks; Geometric Modelling, Graphics and Visualisation; Advanced and Emerging Applications; Information Systems and Technologies; and Urban and Regional Planning.

For these Workshops Proceedings, we accepted 328 full papers, 23 short papers, and 3 Ph.D. Showcase papers from a total of 1184 submissions (acceptance rate 29.9%). Reviews were single blind, and submissions received 2.6 reviews each, on average. In the Main Conference Proceedings, we accepted 119 full papers, 3 short papers, and 2 Ph.D. Showcase papers from 431 submissions to the General Tracks of the Conference (acceptance rate of 29.9%). We would like to convey our sincere appreciation to the workshop chairs and co-chairs and program committee members, for their diligent work, commitment, and dedication.

The success and consistent maintenance of the ICCSA conference series in general, and of ICCSA 2026 in particular, rely upon the support of many people: authors, presenters, participants, keynote speakers, workshop chairs, session chairs, Organizing Committee members, student volunteers, Program Committee members, Advisory Committee members, International Liaison chairs, reviewers, and other individuals in various roles. Thus, we take this opportunity to wholeheartedly thank each and every one.

We also wish to thank the publisher, Springer, for accepting the proceedings, and for their kind assistance and cooperation during the editing process. We would cordially like to invite you to refer to the ICCSA website https://iccsa.org, where you can find the relevant details regarding this academic endeavour and our event.

June 2026

Chiara Garau
Ana Maria A. C. Rocha
Osvaldo Gervasi
Beniamino Murgante

Welcome Message from the Organizers

On behalf of the Local Organizing Committee of ICCSA 2026, it is our great pleasure to welcome you to the 26th International Conference on Computational Science and Its Applications, held from June 30 to July 3, 2026, at the University of Minho in Braga, Portugal.

We are deeply honoured by the trust placed in us by the ICCSA General Chairs in organizing another edition of this successful conference series. We hope that ICCSA 2026 provided an inspiring environment for scientific exchange, collaboration, and discussion among researchers from all over the world.

The conference took place at the University of Minho, one of the most prestigious higher education institutions in Portugal. Founded in 1973, the University of Minho has campuses in Braga and Guimaraes and is widely recognized for its academic excellence, research activity, and strong international outlook.

Braga, one of the oldest cities in Portugal, is renowned for its historical heritage, cultural vitality, and welcoming atmosphere.

Located in the heart of the Minho region, the city combines tradition with modernity, offering visitors renowned gastronomy, beautiful landscapes, and the unique character of *Vinho Verde* wine.

The scientific program of ICCSA 2026 included plenary lectures delivered by distinguished scientists, together with a broad range of workshops covering several areas of Computational Science and its Applications. We believe that the conference provided excellent opportunities to share ideas, discuss emerging challenges, and foster new collaborations.

In addition to the scientific sessions, participants had the opportunity to enjoy a rich social program.

Planned activities included a Welcome Reception at Largo do Paço, the Rectorate of the University of Minho, located in the former Archbishop's Palace - one of the most emblematic historical buildings in the city center; a *Vinho Verde* wine-tasting experience at Quinta d'Amares, where particiapnts could discover the art of winemaking; and a guided visit to the coastal city of Viana do Castelo, renowned for its maritime heritage, traditional architecture, and beautiful Atlantic scenery. The conference gala dinner took place at Colunata Eventos, located within the Bom Jesus do Monte sanctuary, a UNESCO World Heritage Site, where ICCSA participants enjoyed delicious recipes and a wonderful view over the city of Braga.

The success of ICCSA 2026 would not have been possible without the dedication and enthusiasm of many collaborators and volunteers, whose hard work and commitment greatly contributed to the organization of the conference.

On behalf of the Local Organizing Committee, it was our honour to warmly welcome you to the beautiful city of Braga and we hope you had an enjoyable and productive conference.

We enjoyed meeting you in Braga.

Sincerely yours,

On behalf of the Local Organizing Committee.

June 2026 Ana Maria A. C. Rocha

Organization

Honorary General Chairs

Bernady O. Apduhan	Kyushu Sangyo University, Japan
Kenneth C. J. Tan	Sardina Systems, UK

General Chairs

Ana Maria A. C. Rocha	University of Minho, Portugal
David Taniar	Monash University, Australia
Osvaldo Gervasi	University of Perugia, Italy

Program Committee Chairs

Beniamino Murgante	University of Basilicata, Italy
Chiara Garau	University of Cagliari, Italy

International Advisory Committee

Jemal Abawajy	Deakin University, Australia
Dharma P. Agarwal	University of Cincinnati, USA
Jagannath Aryal	University of Melbourne, Australia
Rajkumar Buyya	University of Melbourne, Australia
Claudia Bauzer Medeiros	University of Campinas, Brazil
Carl K. Chang	Iowa State University, USA
Elizabeth Chang	University of the Sunshine Coast, Australia
Manfred M. Fisher	Vienna University of Economics and Business, Austria
Pierre Frankhauser	University of Franche-Comté/CNRS, France
Marina L. Gavrilova	University of Calgary, Canada
Sumi Helal	University of Florida, USA & Lancaster University, UK
Bin Jiang	University of Gävle, Sweden
Yee Leung	Chinese University of Hong Kong, China
Biswajeet Pradhan	University of Technology Sydney, Australia
My Thai	University of Florida, USA

International Liaison Chairs

Giuseppe Borruso	University of Trieste, Italy
Elise De Donker	Western Michigan University, USA
Maria Noelia Faginas Lago	University of Perugia, Italy
Maria Irene Falcão	University of Minho, Portugal
Robert C. H. Hsu	Chung Hua University, Taiwan
Yeliz Karaca	UMass Chan Medical School, USA
Tae-Hoon Kim	Zhejiang University of Science and Technology, China
Vladimir Korkhov	Saint Petersburg University, Russia
Takashi Naka	Kyushu Sangyo University, Japan
Rafael D. C. Santos	National Institute for Space Research, Brazil
Maribel Yasmina Santos	University of Minho, Portugal
Anastasia Stratigea	National Technical University of Athens, Greece
Nguyen Canh Thai	Thuyloi University, Vietnam
Kh Md Nahiduzzaman	Mohammed VI Polytechnic University, Rabat, Morocco

Workshop and Session Organizing Chairs

Beniamino Murgante	University of Basilicata, Italy
Chiara Garau	University of Cagliari, Italy

Award Chair

Wenny Rahayu	La Trobe University, Australia

Publicity Committee Chairs

Elmer Dadios	De La Salle University, Philippines
Nataliia Kulabukhova	Saint Petersburg University, Russia
Daisuke Takahashi	Tsukuba University, Japan
Doan Quang Tu	Academy of Military Science and Technology, Vietnam
Shangwang Wang	Beijing University of Posts and Telecommunications, China

Local Organizing Committee Chairs

Maria Irene Falcão	University of Minho, Portugal
Ana Cristina Braga	University of Minho, Portugal
Rui José	University of Minho, Portugal
José Manuel Machado	University of Minho, Portugal
Jorge G. Rocha	University of Minho, Portugal

Technology Chair

Damiano Perri	University of Perugia, Italy

Workshops

Workshop on Advancements in Applied Machine-learning and Data Analytics (AAMDA 2026)

Workshop Organizers

Alessandro Costantini	INFN, Italy
Elisabetta Ronchieri	INFN CNAF, Italy

Workshop Program Committee Members

Luca Dell'Agnello	INFN, Italy
Barbara Martelli	INFN, Italy
Daniele Cesini	INFN CNAF, Italy
Giovanni Zurlo	INFN CNAF, Italy
Luca Giommi	INFN CNAF, Italy
Valentino Santucci	Università per Stranieri di Perugia, Italy
Francesco Santini	University of Perugia, Italy
Paolo Mengoni	Hong Kong Baptist University, China
Alessandro Bombini	Istituto Nazionale di Fisica Nucleare, Italy

Artificial Intelligence Tools In Managing Future Communities (AiMFutCOM 2026)

Workshop Organizers

Marco Vona	University of Basilicata, Italy
Francesco Scorza	University of Basilicata, Italy
Donatello Cardone	University of Basilicata, Italy
Maurizio Pollino	ENEA, Italy
Gregorio DAgostino	ENEA, Italy

Workshop Program Committee Members

Maria Danese	CNRISPC, Italy
Rachele Vanessa Gatto	Università della Basilicata, Italy
Simone Corrado	University of Basilicata, Italy
Alessandro Plaisant	University of Sassari, Italy
Sabrina Lai	University of Cagliari, Italy
Alfonso Annunziata	University of Basilicata, Italy
Jolanta Dvarioniene	Kaunas University of Technology, Lithuania

Advanced and Innovative Web Apps 2026 (AIWA 2026)

Workshop Organizers

Damiano Perri	University of Perugia, Italy
Osvaldo Gervasi	University of Perugia, Italy
Luciano Alessandro Ipsaro Palesi	University of Florence, Italy

Workshop Program Committee Members

David Berti	ART SpA, Italy
JungYoon Kim	Gachon University, South Korea
Taihoon Kim	Chonnam National University, South Korea

Advanced Mathematics and Computing Models in Complex Data-Intensive Computational Systems (AMCM 2026)

Workshop Organizers

Yeliz Karaca	UMass Chan Medical School and Massachusetts Institute of Technology, USA
Dumitru Baleanu	Lebanese American University, Lebanon
Osvaldo Gervasi	University of Perugia, Italy
Yudong Zhang	University of Leicester, UK
Majaz Moonis	UMass Chan Medical School, USA

Workshop Program Committee Members

Aziz Dursun	Virginia Tech University, USA
Akif Akgul	Hitit University, Turkey
Nizamettin Aydin	Istanbul Technical University, Turkey
Martin Bohner	Missouri University of Science and Technology, USA
Khan Muhammad	Sungkyunkwan University, South Korea

Advanced Numerical Approaches for assessment and design of no-tension Masonry Structures (ANAMS 2026)

Workshop Organizers

Antonino Iannuzzo	Universitá degli studi del Sannio, Italy
Natalia Pingaro	Politecnico di Milano, Italy
Andrea Montanino	International Centre for Numerical Methods in Engineering, Spain
Elham Mousavian	University of Edinburgh, UK
Stefano Galassi	University of Florence, Italy

Workshop Program Committee Members

Amal Gerges	University of Cagliari, Italy
Vittorio Paris	University of Bergamo, Italy
Alessandra Capolupo	Polytechnic University of Bari, Italy
Carlo Olivieri	Pegaso University, Italy
Generoso Vaiano	Pegaso University, Italy

Luigi Sibille	Norwegian University of Science and Technology, Norway
Francesco Picariello	Universitas Mercatorum, Italy
Ilaria Scarcelli	Politecnico di Bari, Italy
Fabian Bernal Orozco Barrera	National Autonomous University of Mexico, Mexico
Gabriel Barbat	Universitat Politècnica de Catalunya, Spain
Alejandro Cornejo	Universitat Politècnica de Catalunya, Spain
Mohammad Pourfouladi	Politecnico di Milano, Italy

Artificialization of land, Urban Transformation, and Territorial Resilience in a changing Climate (AUTTREC 2026)

Workshop Organizers

Celestina Fazia	Università di Enna Kore, Italy
Kh Md Nahiduzzaman	Mohammed VI Polytechnic University, Morocco
Hassan Radoine	Mohammed VI Polytechnic University, Morocco
Cezar Morar	University of Oradea, Romania
Antonio García Martínez	Universidad de Sevilla, Spain
Clara Stella Vicari Aversa	Università degli Studi Mediterranea di Reggio Calabria, Italy

Workshop Program Committee Members

Federica Sortino	Universitià degli studi di Enna Kore, Italy
Giulia Fernanda Grazia Catania	Universit di Enna Kore, Italy
Maria Francesca Faro	Ministry of Education and Merit (MIM), Italy
Aleksandar Valjarević	University of Belgrade, Serbia
Laila El Ghazouani	Mohammed VI Polytechnic University, Morocco
Adnane Labbaci	Ibn Tofail University, Morocco

Blockchain and Distributed Ledgers: Technologies and Applications (BDLTA 2026)

Workshop Organizers

Vladimir Korkhov	Saint Petersburg State University, Russia
Elena Stankova	Saint Petersburg State University, Russia
Nataliia Kulabukhova	AutoTech, Russia

Workshop Program Committee Members

Vasily Solodkov	National Research University Higher School of Economics, Moscow, Russia
Alexander Bogdanov	St. Petersburg State University, Russia
Alexander Degtyarev	Saint Petersburg State University, Russia
Dmitrii Vasiunin	Deutsche Telekom Cloud Services E.P.E., Greece
Serob Balyan	Institute for Informatics and Automation Problems, National Academy of Sciences of Armenia, Armenia
Suren Abrahamyan	Institute for Informatics and Automation Problems of NAS RA, Armenia
Ashot Gevorkyan	Institute for Informatics and Automation Problems NAS of Armenia, Armenia
Martin Vala	Pavol Jozef Šafárik University, Slovak Republic
Nodir Zaynalov	Samarkand Branch of Tashkent University of Information Technologies Named After Muhammad al Khwarizmi, Uzbekistan
Michail Panteleyev	Saint Petersburg Electrotechnical University, Russia
Alexander Degtyarev	St. Petersburg State University, Russia
Alexander Bogdanov	St. Petersburg State University, Russia

Bio and Neuro inspired Computing and Applications (BIONCA 2026)

Workshop Organizers

Nadia Nedjah	State University of Rio de Janeiro, Brazil
Luiza de Macedo Mourelle	State University de Rio de Janeiro, Brazil

Workshop Program Committee Members

Alan Oliveira	Universidade de Lisboa, Portugal
Rogerio Calazan	Instituto de Estudos do Mar Almirante Paulo Moreira - Brazilian Navy, Brazil
Marcos Farias	Instituto de Engenharia Nuclear, Brazil
Joelmir Ramos	State University de Rio de Janeiro, Brazil
Luigi Maciel Ribeiro	State University of Rio de Janeiro, Brazil

Computational and Applied Mathematics (CAM 2026)

Workshop Organizers

Maria Irene Falcao	University of Minho, Portugal
Fernando Miranda	Universidade do Minho, Portugal

Workshop Program Committee Members

Graca Tomaz	Instituto Politécnico da Guarda, Portugal
Helmuth Malonek	University of Aveiro, Portugal
Isabel Cacao	University of Aveiro, Portugal
Joao Morais	Autonomous Technological Institute of Mexico (ITAM), Mexico
Lidia Aceto	Università del Piemonte Orientale, Italy
Luís Ferrás	University of Minho, Portugal
Patricia Beites	University of Beira Interior, Portugal
Paulo Amorim	Escola de Matemática Aplicada da Fundação Getulio Vargas, Brazil
Regina de Almeida	University of Trás-os-Montes and Alto Douro, Portugal
Ricardo Severino	University of Minho, Portugal

Computational and Applied Statistics for Data & AI (CAS 2026)

Workshop Organizers

Ana Cristina Braga	University of Minho, Portugal
Ines Sena	University of Minho, Portugal
Joao Mendes	Polytechnic Institute of Bragança, Portugal

Workshop Program Committee Members

A. Manuela Goncalves	University of Minho, Portugal
Vera Afreixo	University of Aveiro, Portugal
Ivan Rodriguez Conde	University of Arkansas at Little Rock, USA
Isabel Natário	Universidade Nova de Lisboa, Portugal
Luca Parodi	University of Genoa, Italy
Flávia Pires	Instituto Politecnico de Bragança, Portugal
Beatriz Flamia Azevedo	Instituto Politecnico de Bragança, Portugal
Adriano Silva	Instituto Politecnico de Bragança VAT 600013758, Portugal

Ana Paula Amorim	Universidade do Minho, Portugal
Angela Silva	Universidade do Minho, Portugal
Arunasalam Rahunanthan	Central State University, USA
Bruna Ramos	University of Minho, Portugal
Cecília Maria Vasconcelos Costa e Castro	Universidade do Minho, Portugal
Florbela Correia	Polytechnic Institute of Viana do Castelo, Portugal
Joaquim Goncalves	Polytechnic Institute of Cávado and Ave, Portugal
Lino Costa	University of Minho, Portugal
Luís Miguel Matos	Universidade do Minho, Portugal
Marco Reis	University of Coimbra, Portugal
Maria Mourao	Polytechnic Institute of Viana do Castelo, Portugal
Martin Perez Perez	University of Vigo, Spain
Sandra Aleixo	FCiências.ID, Portugal
Paula Odete Fernandes	Instituto Politecnico de Bragança, Portugal
Iván Rodríguez Conde	Universidade de Vigo, Spain
Telmo Miguel Pires Pinto	University of Minho, Portugal

Circular Economy and Territorial Competitiveness (CETC 2026)

Workshop Organizers

Gerardo Carpentieri	University of Naples Federico II, Italy
Carmen Guida	Università degli Studi di Napoli Federico II, Italy
Floriana Zucaro	University of Naples Federico II, Italy
Zhadyra Konurbaeva	D. Serikbayev East Kazakhstan Technical University, Kazakhstan
Saule Kumargazhanova	D. Serikbayev East Kazakhstan Technical University, Kazakhstan

Workshop Program Committee Members

Rosaria Battarra	ISMed-CNR, Italy
Laura Ascione	Università degli studi di Napoli Federico II, Italy
Valerio Martinelli	Università di Napoli Federico II, Italy
Madina Yussubaliyeva	D. Serikbayeva East Kazakhstan Technical University, Kazakhstan
Samat Baigereyev	D. Serikbayev East Kazakhstan Technical University, Kazakhstan

Circular Ecological Heritage Communities: Regeneration and Evaluation Strategies for Historic Villages, Landscapes and Rural Territories (CHEC 2026)

Workshop Organizers

Mariarosaria Angrisano	Pegaso University, Italy
Antonia Gravagnuolo	CNR Institute of Heritage Science, Italy
Francesca Buglione	CNR ISPC, Italy
Piera Della Morte	Pegaso Telematic University, Italy

Workshop Program Committee Members

Giulia Datola	Politecnico di Milano, Italy
Marco Rossitti	Politecnico di Milano, Italy
Federico DellAnna	Politecnico di Torino, Italy
Grazia Neglia	Pegaso Telematic University, Italy
Vanessa Assumma	University of Bologna, Italy
Ippolita Mecca	Pegaso Telematic University, Italy
Ferdinando Verardi	Pegaso Telematic University, Italy
Eleonora Sottile	University of Cagliari, Italy
Dina Jovanovic	Politecnico di Milano, Italy
Martina Mulinacci	Politecnico di Milano, Italy
Davide Cerati	Politecnico di Milano, Italy
Daniela Oreni	Politecnico di Milano, Italy
Patrizia Riganti	University of Glasgow, UK
Carlo Olivieri	Pegaso University, Italy
Giulio Cavana	Politecnico di Torino, Italy
Maria Tartari	CNR, Italy
Luigi Fusco Girard	Pegaso Telematic University, Italy
Albina Scioti	Università Telematica Pegaso, Italy
Eleonora Righetto	Università degli Studi di Padova, Italy
Chiara D'Alpaos	University of Padua, Italy
Anthony Suppa	Politecnico di Torino, Italy
Giorgia Malavasi	Politecnico di Torino, Italy
Kaoutare Amini Alaoui	Mohamed VI Polytechnic University, Morocco
Elena Todella	University of Parma, Italy
Leonhard Seyboth	Federal University of Santa Catarina, Brazil
Jianzhuo Xu	Politecnico di Milano, Italy
Ludovica Nasca	University of Catania, Italy
Alessandro Plaisant	University of Sassari, Italy

Maria Carla Saliu	University of Cagliari, Italy
Ons Zoghlami	Sapienza University of Rome, Italy
Irena Ištoka Otković	Josip Juraj Strossmayer University of Osijek, Croatia
Daniele Cannatella	TU Delft, Italy
Tanja Congiu	University of Sassari, Italy
Juan Camilo Giraldo	Institute for Advanced Architecture of Catalonia, Spain
YiChang Liu	Bio-Architecture Formosana, Taiwan
Yasmine Selim	Ain Shams University, Egypt
Simona Tondelli	Alma Mater Studiorum University of Bologna, Italy
Efrossyni Tsakiri	National Technical University of Athens, Greece

Cyber Intelligence and Applications (CIA 2026)

Workshop Organizers

Gianni D'Angelo	University of Salerno, Italy
Francesco Palmieri	University of Salerno, Italy
Massimo Ficco	University of Salerno, Italy

Workshop Program Committee Members

Arcangelo Castiglione	University of Salerno, Italy
Ugo Fiore	University of Salerno, Italy
Michele Mastroianni	University of Foggia, Italy
Marcello Trovati	Edge Hill University, UK
Raffaele Esposito	University of Salerno, Italy
Marek Ogiela	AGH University of Science and Technology, Poland
Chang Choi	Gachon University, South, South Korea
Mauro Migliardi	University of Padua, Italy

Computational Methods, Statistics and Industrial Mathematics (CMSIM 2026)

Workshop Organizers

Maria Filomena Teodoro	CINAV, Escola Naval, Instituto Universitário Militar, Portugal

Paula Simões	Academia Militar, Portugal
Teresa A. Oliveira	University of Lisbon and Universidade Aberta, Portugal

Workshop Program Committee Members

Luis Teodoro	University of Oslo, Norway
Eliana Costa e Silva	Polytechnic of Porto, Portugal
Ana Borges	Instituto Politécnico do Porto, Portugal
Cristina Lopes	Polytechnic Institute of Porto, Portugal
Fernando Lucas Carapau	University of Évora, Portugal
Antonio S. Macias	University of A Coruña, Spain
Matteo Sacchet	University of Turin, Italy
Marina Marchisio Conte	University of Turin, Italy
Christos Kitsos	University of West Attica, Greece
M. Rosário Ramos	Universidade Aberta, University of Lisbon, Portugal
António Pacheco	Instituto Superior Técnico, Portugal
Ana Sofia Rézio	Polytechnic Institute of Lisbon, Portugal
Paula Simões	Military Academy Portuguese Army, Portugal
Ricardo Moura	NOVA University of Lisbon, Portugal
Ana Oliveira	Academia Militar, Portugal
Sílvia Chá	Military Academy, Portugal
Carlos Rodrigues	Instituto Politecnico de Bragança, Portugal

Computational Optimization and Applications (COA 2026)

Workshop Organizers

Ana Maria A. C. Rocha	University of Minho, Portugal
Humberto Rocha	University of Coimbra, Portugal

Workshop Program Committee Members

Eligius Hendrix	Universidad de Málaga, Spain
Joana Dias	University of Coimbra, Portugal
Ana Pereira	Polytechnic Institute of Bragança, Portugal
Lino Costa	University of Minho, Portugal, Portugal
Filipe Alvelos	Universidade do Minho, Portugal
Ana Paula Teixeira	Universidade do Minho, Portugal

Florbela P. Fernandes	Politecnico de Bragança, Portugal
Ana Monteiro	University of Coimbra, Portugal
Rui Pascoal	University of Coimbra, Portugal
Rommel Regis	Saint Joseph's University, USA
Telmo Miguel Pires Pinto	University of Minho, Portugal
Emerson Paiva	Minho University, Portugal

Computational Methods in Advanced Networking Technologies (CoMANT 2026)

Workshop Organizers

Eslam Farsimadan	University of Salerno, Italy
Leila Moradi	University of Salerno, Italy
Francesco Palmieri	University of Salerno, Italy

Workshop Program Committee Members

Massimo Ficco	University of Salerno, Italy
Ugo Fiore	University of Salerno, Italy
Michele Mastroianni	University of Foggia, Italy
Arcangelo Castiglione	University of Salerno, Italy
Francesco Giuseppe Zammarrelli	University of Salerno, Italy
Annalaura Miglino	University of Salerno, Italy
Mojtaba Sharifzadeh	Università di Trento, Italy
Zahra Ebadi Ansaroudi	Fondazione Bruno Kessler, Italy
Mohammad Vazir Panah	University of Salerno, Italy
Gennaro Pio Rimoli	Università degli Studi di Salerno, Italy

Computational Astrochemistry 2026 (CompAstro 2026)

Workshop Organizers

Marzio Rosi	University of Perugia, Italy
Daniela Ascenzi	Università degli Studi di Trento, Italy
Nadia Balucani	University of Perugia, Italy
Stefano Falcinelli	University of Perugia, Italy

Workshop Program Committee Members

Luca Mancini	University of Perugia, Italy
Andrea Giustini	University of Perugia, Italy
Albert Rimola	Universitat Autònoma de Barcelona, Spain
Piero Ugliengo	Università degli Studi di Torino, Italy
Dario Campisi	University of Perugia, Italy
Gianmarco Valuzzo	University of Perugia, Italy
Giacomo Pannacci	Università degli Studi di Perugia, Italy
Giacomo Giorgi	Università degli Studi di Perugia, Italy
Gabriella Di Genova	Università degli Studi di Perugia, Italy
Lisa Giani	Université Grenoble Alpes, IPAG, France
Emilia Valenca Ferreira de Aragao	University of Montpellier, CEA, CNRS, ENSCM, France
Marco Parriani	Università degli studi di Perugia, Italy
Costanza Borghesi	CNR Istituto Nanoscienze, Italy

Computational methods for Porous materials (CompPor 2026)

Workshop Organizers

Vadim Lisitsa	Institute of Petroleum Geology and Geophysics SB RAS, Russia
Evgeniy Romenski	Sobolev Institute of Mathematics, Russia

Workshop Program Committee Members

Tatyana Khachkova	Institute of Petroleum Geology and Geophysics SB RAS, Russia
Dmitry Prokhorov	Institute of Petroleum Geology and Geophysics SB RAS, Russia
Mikhail Novikov	Sobolev Institute of Mathematics SB RAS, Russia
Galina Reshetova	Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Russia
Sergey Solovyev	Sobolev Institute of Mathematics SB RAS, Russia
Vladimir Cheverda	Sobolev Institute of Mathematics SB RAS, Russia

Workshop on Computational Science and HPC (CSHPC 2026)

Workshop Organizers

Elise de Doncker	Western Michigan University, USA
Hideo Matsufuru	High Energy Accelerator Research Organization, Japan
Robert Makin	Western Michigan University, USA

Workshop Program Committee Members

Hiroshi Daisaka	Hitotsubashi University, Japan
James Rhodes	Western Michigan University, USA
Naohito Nakasato	University of Aizu, Japan
Issaku Kanamori	RIKEN, Japan
Takahiro Ueda	Juntendo University, Japan
Keigo Nitadori	RIKEN, Japan
Fukuko Yuasa	KEK, Japan
Shigeyoshi Aoyama	KEK, Japan
Tatsumi Aoyama	University of Tokyo, Japan
Tasnim Gharaibeh	Kalamazoo College, USA

Cities, Technologies and Planning (CTP 2026)

Workshop Organizers

Giuseppe Borruso	University of Trieste, Italy
Beniamino Murgante	University of Basilicata, Italy
Malgorzata Hanzl	Lodz University of Technology, Poland
Anastasia Stratigea	National Technical University of Athens, Greece
Ljiljana Zivkovic	Republic Geodetic Authority, Serbia
Ginevra Balletto	University of Cagliari, Italy

Workshop Program Committee Members

Silvia Battino	University of Sassari, Italy
Mara Ladu	University of Cagliari, Italy
Maria del Mar Munoz	University of Cádiz, Spain
Ainhoa Amaro	EVOCONS, Spain
Enrico Dagostini	University of Malta, Malta

Francesca Krasna	Università degli Studi di Trieste, Italy
Brisol García García	Universidad Politécnica de Quintana Roo, Mexico
Giovanni Mauro	University of Campania Luigi Vanvitelli, Italy
Maria Ronza	University of Naples Federico II, Italy
Clara Di Fazio	University of Trieste, Italy
Andrea Gallo	Università degli Studi di Trieste, Italy
Francesca Sinatra	University of Trieste, Italy

Evaluating Inner Areas Potentials (EIAP 2026)

Workshop Organizers

Diana Rolando	Politecnico di Torino, Italy
Manuela Rebaudengo	Politecnico di Torino, Italy
Alice Barreca	Politecnico di Torino, Italy
Giorgia Malavasi	Politecnico di Torino, Italy

Workshop Program Committee Members

John Accordino	Virginia Commonwealth University, USA
Sara Favargiotti	University of Trento, Italy
Barbara Lino	Università degli Studi di Palermo, Italy
Maddalena Ferretti	Università Politecnica delle Marche, Italy
Eugenio Muccio	University of Naples Federico II, Italy
Giuliano Poli	University of Naples Federico II, Italy
Giulia Datola	Politecnico di Milano, Italy
Elena Todella	University of Parma, Italy
Asja Aulisio	Politecnico di Torino, Italy
Marco Rossitti	Politecnico di Milano, Italy
Martina Spinelli	Politecnico Torino, Italy
Amina Pereno	Politecnico di Torino, Italy
Francesco Bruzzone	Università Iuav di Venezia, Italy
Alexandra Stankulova	Politecnico di Torino, Italy
Benedetta Di Leo	Università Politecnica delle Marche, Italy
Edda Donati	University of Rome Tor Vergata, Italy
Vanessa Assumma	University of Bologna, Italy
Beatrice Mecca	Politecnico di Torino, Italy

sustainablE mobiLity and Last mIle lOgisTic in urban context (ELLIOT 2026)

Workshop Organizers

Tiziana Campisi	University of Enna Kore, Italy
Socrates Basbas	Aristotle University of Thessaloniki, Greece
Florin Rusca	National University of Science and Technology Politehnica Bucharest, Romania
Grigorios Fountas	Aristotle University of Thessaloniki, Greece
Aleksandra Deluka-Tibljas	University of Rijeka, Croatia

Workshop Program Committee Members

Francesco Russo	University of Reggio Calabria, Italy
Oana Dinu	National University of Science and Technology Politehnica Bucharest, Romania
Antonio Polimeni	Università di Messina, Italy
Antonio Russo	University of Enna Kore, Italy
Elena Cocuzza	Univesity of Catania, Italy
Corrado Rindone	Università degli studi Mediterranea di Reggio Calabria, Italy

Econometrics and Multidimensional Evaluation in Urban Environment (EMEUE 2026)

Workshop Organizers

Maria Cerreta	University of Naples Federico II, Italy
Giuseppe Ciciriello	University of Naples Federico II, Italy
Simona Panaro	University of Sussex Business School, UK
Piero Zizzania	Università of Naples Federico II, Italy
Pierluigi Morano	Polytechnic University of Bari, Italy
Francesco Tajani	Sapienza University of Rome, Italy

Workshop Program Committee Members

Giuliano Poli	Università Federico II di Napoli, Italy
Ludovica La Rocca	University of Naples Federico II, Italy
Sabrina Sacco	University of Naples Federico II, Italy

Ecosystem Services in Spacial Planning for climate neutral urban and rural areas (ESSP 2026)

Workshop Organizers

Sabrina Lai	University of Cagliari, Italy
Carmela Gargiulo	Università Federico II di Napoli, Italy
Beniamino Murgante	University of Basilicata, Italy
Francesco Scorza	University of Basilicata, Italy
Corrado Zoppi	University of Cagliari, Italy
Floriana Zucaro	University of Naples Federico II, Italy

Workshop Program Committee Members

Alfonso Annunziata	University of Basilicata, Italy
Ginevra Balletto	University of Cagliari, Italy
Ivan Blečić	University of Cagliari, Italy
Giuseppe Borruso	University of Trieste, Italy
Barbara Caselli	Università degli Studi di Parma, Italy
Maria Cerreta	University of Naples Federico II, Italy
Carmen Guida	Università degli Studi di Napoli Federico II, Italy
Federica Isola	University of Cagliari, Italy
Francesca Leccis	University of Cagliari, Italy
Federica Leone	University of Cagliari, Italy
Silvia Rossetti	Università degli Studi di Parma, Italy
Luigi Santopietro	Institute of Methodologies for Environmental Analysis CNR IMAA, Italy
Simone Corrado	University of Basilicata, Italy
Rachele Vanessa Gatto	Università della Basilicata, Italy

The 16th International Workshop on Future Information System Technologies and Applications (FiSTA 2026)

Workshop Organizers

Bernady O. Apduhan	Kyushu Sangyo University, Japan
Rafael Santos	National Institute for Space Research (INPE), Brazil

Workshop Program Committee Members

Agustinus Waluyo	La Trobe University, Australia
Earl Ryan Aleluya	Mindanao State University Iligan Institute of Technology, Philippines
Eric Pardede	La Trobe University, Australia
Kai Cheng	Kyushu Sangyo University, Japan
Fenghui Yao	Tennessee State University, USA
Yusuke Gotoh	Okayama University, Japan
Kazuaki Tanaka	Kyushu Institute of Technology, Japan
Toshihiro Yamauchi	Okayama University, Japan
Toshihiro Uchibayashi	Kyushu University, Japan
Krishnamoorthy Ranganathan	Chennai Institute of Technology, India
Orven Llantos	MSU-Iligan Institute of Technology, Philippines
Anton Louise De Ocampo	Batangas State University, Philippines

Geographical Analysis, Urban Modeling, Spatial Statistics (Geog-And-Mod 2026)

Workshop Organizers

Giuseppe Borruso	University of Trieste, Italy
Beniamino Murgante	University of Basilicata, Italy
Hartmut Asche	Hasso Plattner Institute for Digital Engineering, Germany
Rodrigo Tapia-McClung	Centro de Investigación en Ciencias de Información Geoespacial, Mexico
Andreas Fricke	Hasso Plattner Institute for Digital Engineering, University of Potsdam, Germany

Workshop Program Committee Members

Ginevra Balletto	University of Cagliari, Italy
Mara Ladu	University of Cagliari, Italy
Gabriella Schoier	Università di Trieste, Italy
Marco Mazzarino	Iuav University of Venice, Italy
Malgorzata Hanzl	Lodz University of Technology, Poland
Anastasia Stratigea	National Technical University of Athens, Greece
Maria del Mar Munoz	University of Cádiz, Spain
Veronica Camerada	Università degli Studi di Sassari, Italy

Generative and Neuro-Symbolic Artificial Intelligence for Real-World Systems (GNS-AI 2026)

Workshop Organizers

Hugo Peixoto	University of Minho, Portugal
Manuel Rodrigues	University of Minho, Portugal
Regina Sousa	University of Minho, Portugal

Workshop Program Committee Members

Dalila Duraes	Universidade do Minho, Portugal
Paulo Novais	Universidade do Minho, Portugal
Laura Cavalcante	Instituto Superior de Engenharia do Porto, Portugal
Diogo Martinho	Instituto Superior de Engenharia de Porto, Portugal
Hélia Guerra	University of the Azores, Portugal
Rohit Raja	Guru Ghasidas Vishwavidyalaya, India
Hien Nguyen	University of Information Technology, Vietnam
Antonio Carlos Da Silva Abelha	Universidade do Minho, Portugal
Cristiana Neto	University of Minho, Portugal
José Machado	University of Minho, Portugal
Luís Conceicao	Instituto Superior de Engenharia do Porto, Portugal
Omar Gomez	Universidad Nacional de Colombia, Colombia
Luis Mendes Gomes	University of the Azores, Portugal
Nicolas Lori	University of the Azores, Portugal
Sara Balderas Díaz	University of Cádiz, Spain
Gabriel Guerrero Contreras	University of Cádiz, Spain
Francisco Zayas Gato	Universidade da Coruña, Spain
Diana Ferreira	University of Minho, Portugal
Jolanta Miliauskaite	Vilnius University, Lithuania

Geomatics for Resource Monitoring and Management (GRMM 2026)

Workshop Organizers

Alberico Sonnessa	Politecnico di Bari, Italy
Eufemia Tarantino	Politecnico di Bari, Italy
Marco Locurcio	Polytechnic University of Bari, Italy

Workshop Program Committee Members

Valeria Belloni	Sapienza University of Rome, Italy
Roberta Ravanelli	University of Liège, Belgium
Michela Ravanelli	Sapienza Università di Roma, Italy
Alessandra Mascitelli	University of Chieti Pescara, Italy
Francesco Bimbo	University of Foggia, Italy
Francesco Di Capua	University of Basilicata, Italy
Gaetano Falcone	Università degli Studi di Napoli Federico II, Italy
Francesco Tajani	Sapienza University of Rome, Italy
Paola Amoruso	LUM Giuseppe Degennaro University, Italy
Pierfrancesco De Paola	University of Naples Federico II, Italy
Isidora Đurić	University of Novi Sad, Serbia
Rossella Petti	Politecnico di Bari, Italy
Valeria Cera	University of Naples Federico II, Italy
Antonino Iannuzzo	Universitá degli studi del Sannio, Italy
Pier Francesco Recchi	Università di Napoli Federico II, Italy
Giuseppe Francesco Cesare Lama	University of Naples Federico II, Italy
Ester Scotto di Perta	University of Naples Federico II, Italy
Antonino Maltese	University of Palermo, Italy
Andrea Montanino	University of Naples, Italy
Carla Cavallo	University of Salerno, Italy
Dario Esposito	Polytechnic University of Bari, Italy
Rosaria Parente	Universitas Mercatorum, Italy
Claudia Panciera	Water Research Institute (IRSA) CNR, Italy
Pasquale Perrini	Consorzio Interuniversitario per l'Idrologia, Italy
Beniamino Onorati	Università degli Studi della Basilicata, Italy
Francesco Sica	Sapienza University of Rome, Italy

International Workshop on Information and Knowledge in the Internet of Things (IKIT 2026)

Workshop Organizer

Teresa Guarda	Universidad Estatal Península de Santa Elena, Ecuador

Workshop Program Committee Members

Isabel Lopes	Instituto Politécnico de Bragança, Portugal
Luis Mazon	Bitrum Research Group, Spain

Arnulfo Alanis	Tecnológico Nacional de México, Campus Tijuana, Mexico
Filipe Mota Pinto	Instituto Politécnico de Leiria, Portugal
Luis Enrique Chuquimarca Jiménez	Universidad Estatal Penìnsula de Santa Elena, Ecuador
Gustavo Gatica	Universidad Andres Bello, Chile
Jairo Coronado Hernandez	Universidad de la Costa, Colombia
Maria Fernanda Augusto	BITrum Research Center, Spain
José María Díaz Nafría	Universidad a Distancia de Madrid, Spain
Modestos Stavrakis	University of the Aegean, Greece
Isabel Ribeiro	Instituto Politécnico Bragança, Portugal
Bruno Sousa	University of Coimbra, Portugal
Datzania Villao	Universidad Estatal Península de Santa Elena, Ecuador

Intelligent Sustainable Development: from Models to Real-World Impact (ISD 2026)

Workshop Organizers

Diego Brandão	Federal Center for Technological Education of Rio de Janeiro, Brazil
Angela Ferreira	Polytechnic University of Bragança, Portugal
Eduardo Bezerra	Federal Center for Technological Education of Rio de Janeiro, Brazil
Raphael Guerra	Universidade Federal Fluminense, Brazil
Douglas Cardoso	University of Coimbra, Portugal
Rui Pedro Lopes	Instituto Politecnico de Bragança, Portugal

Workshop Program Committee Members

Wagner Telles	Universidade Federal Fluminense, Brazil
Jader Lugon Junior	Instituto Federal Fluminense, Brazil
Edcarllos Santos	Universidade Federal do Sul da Bahia, Brazil
Kennedy Fernandes	Universidade Federal de Juiz de Fora, Brazil
Israel Mendonca	Kumamoto University, Japan
Pedro Henrique Gonzalez	Federal University of Rio de Janeiro, Brazil
Pablo Luiz Araujo Munhoz	Federal University of Ouro Preto, Brazil
Elloa Guedes	Amazonas State University, Brazil
Danielli Araújo Lima	Instituto Federal do Triângulo Mineiro, Brazil
Ary Henrique Oliveira	Universidade Federal de Tocantins, Brazil

Paula Odete Fernandes	Instituto Politecnico de Bragança, Portugal
Hiago Rocha	National Laboratory for Scientific Computing, Brazil
Andre Luiz Coutinho Merlo	Centro Federal de Educacao Tecnologica Celso Suckow da Fonseca CEFET RJ, Brazil
Kennedy Fernandes	Universidade Federal de Juiz de Fora, Brazil
Ary Henrique Oliveira	Universidade Federal de Tocantins, Brazil
Mayara Omai	Centro Federal de Educacao Tecnologica Celso Suckow da Fonseca CEFET RJ, Brazil

Multidimensional Evolutionary Evaluations for Transformative Approaches (MEETA 2026)

Workshop Organizers

Maria Cerreta	University of Naples Federico II, Italy
Giuliano Poli	Università Federico II di Napoli, Italy
Gaia Daldanise	CNR Institute for Research on Innovation and Services for Development, Italy
Ludovica La Rocca	University of Federico II of Naples, Italy
Maria Somma	University of Naples Federico II, Italy
Sabrina Sacco	University of Naples Federico II, Italy

Workshop Program Committee Members

Chiara Mazzarella	TUDelft, The Netherlands
Simona Panaro	University of Sussex Business School, UK
Eugenio Muccio	University of Naples Federico II, Italy

Building multi-dimensional Models for assessing complex Environmental Systems (MES 2026)

Workshop Organizers

Vanessa Assumma	University of Bologna, Italy
Caterina Caprioli	Politecnico di Torino, Italy
Giulia Datola	Politecnico di Milano, Italy
Federico DellAnna	Politecnico di Torino, Italy
Marta Dell'Ovo	Politecnico di Milano, Italy
Marco Rossitti	Politecnico di Milano, Italy

Workshop Program Committee Members

Mariarosaria Angrisano	Pegaso University, Italy
Giuliano Poli	University of Naples Federico II, Italy
Francesco Sica	Sapienza University of Rome, Italy
Eugenio Muccio	University of Naples Federico II, Italy
Sebastiano Barbieri	Politecnico di Torino, Italy
Giulio Cavana	Politecnico di Torino, Italy
Marta Bottero	Politecnico di Torino, Italy
Francesca Torrieri	Politecnico di Milano, Italy
Maksims Feofilovs	Riga Technical University, Latvia
Daniela Maiullari	Delft University of Technology, The Netherlands
Elena Todella	University of Parma, Italy
Yvonne Russo	University of Campania L. Vanvitelli, Italy
Tommaso Lazzarin	University of Padua, Italy
Francesco Tajani	Sapienza University of Rome, Italy
Claudia Amadei	University of Padua, Italy
Isabella Maria Lami	Politecnico di Torino, Italy
Federico Vagnon	Politecnico di Torino, Italy
Francesca Torrieri	Politecnico di Milano, Italy
Elena De Angelis	Politecnico di Torino, Italy
Piergiorgio Farfariello	Politecnico di Torino, Italy
Francesca Biondi	Politecnico di Torino, Italy
Giulio Cavana	Politecnico di Torino, Italy
Fabiana Forte	Università della Campania Luigi Vanvitelli, Italy
Siyu Qiu	University of Bologna, Italy
Eugenio Muccio	University of Naples Federico II, Italy
Iuliia Kozlova	University of Bologna, Italy
Danny Casprini	Politecnico di Milano, Italy
Franco Corti	University of Padua, Italy
Elisa Sbardella	Politecnico di Torino, Italy
Andrea Sarcina	Università di Padova, Italy
Daniela Menna	Università degli studi della Campania Luigi Vanvitelli, Italy
Sebastiano Barbieri	Politecnico di Torino, Italy
Ana Soares	University of Minho, Portugal
Francesco Sica	Sapienza University of Rome, Italy
Chiara D'Alpaos	University of Padua, Italy
Alessandra Oppio	Politecnico di Milano, Italy
Maria Rosaria Guarini	Sapienza University of Rome, Italy
Alessia Crisopulli	Politecnico di Milano, Italy
Elisabetta Colucci	Politecnico di Torino, Italy

Giorgia Sugoni	Links Foundation, Italy
Lucia Petronio	University of Bologna, Italy
Marco Alvise Bragadin	University of Bologna, Italy

Computational Mechanics, Intelligence, and Data (MIND 2026)

Workshop Organizers

Lelio Campanile	Università degli Studi della Campania Luigi Vanvitelli, Italy
Renato Zona	Università della Campania Vanvitelli, Italy

Workshop Program Committee Members

Vincenzo Minutolo	Università degli Studi della Campania Luigi Vanvitelli, Italy
Fiammetta Marulli	University of Campania Luigi Vanvitelli, Italy
Simone Palladino	Università degli Studi della Campania Luigi Vanvitelli, Italy
Ali Forooghi	University of Campania Luigi Vanvitelli, Italy
Mario Argenziano	University of Palermo, Italy
Donato Perfetto	University of Campania Luigi Vanvitelli, Italy

Modelling Liveable Cities: Techniques, Methods, Challenges, and Perspectives Behind the X-Minute City (MLC 2026)

Workshop Organizers

Federico Mara	University of Pisa, Italy
Valerio Cutini	University of Pisa, Italy

Workshop Program Committee Members

Beniamino Murgante	University of Basilicata, Italy
Chiara Garau	University of Cagliari, Italy
Francesco Zullo	University of LAquila, Italy
Alessandro Marucci	University of L Aquila, Italy
Chiara Di Dato	University of L Aquila, Italy
Giampiero Lombardini	University of Genoa, Italy

Alfonso Annunziata	University of Basilicata, Italy
Simone Corrado	University of Basilicata, Italy
Müslüm Hacar	Yildiz Technical University, Turkey
Federica Deri	University of Pisa, Italy
Chiara Anselmi	University of Pisa, Italy

Mathematical Methods for Image Processing and Understanding 2026 (MMIPU 2026)

Workshop Organizers

Ivan Gerace	Università degli Studi di Perugia, Italy
Gianluca Vinti	Università degli Studi di Perugia, Italy
Arianna Travaglini	Lucian Blaga University of Sibiu, Romania

Workshop Program Committee Members

Marco Baioletti	University of Perugia, Italy
Marco Donatelli	University of Insubria, Italy
Muhammad Hanif	Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Pakistan
Francesco Marchetti	University of Padua, Italy
Wolfgang Erb	Università degli Studi di Padua, Italy
Ana Maria Acu	Lucian Blaga University of Sibiu, Romania
Simone Rebegoldi	University of Modena and Reggio Emilia, Italy
Valentina Giorgetti	Università degli Studi di Perugia, Italy
Vita Leonessa	University of Basilicata, Italy

MOdels and indicators for assessing and measuring the urban settlement deVElopment in the view of NET ZERO by 2050 (MOVEto0 2026)

Workshop Organizers

Lorena Fiorini	University of L'Aquila, Italy
Lucia Saganeiti	Institute of Methodologies for Environmental Analysis (IMAA) CNR, Italy
Angela Pilogallo	Institute of Methodologies for Environmental Analysis (IMAA) CNR, Italy

Francesco Zullo	University of L'Aquila, Italy
Alessandro Marucci	University of L'Aquila, Italy

Workshop Program Committee Members

Ginevra Balletto	University of Cagliari, Italy
Giuseppe Borruso	University of Trieste, Italy
Chiara Garau	University of Cagliari, Italy
Beniamino Murgante	University of Basilicata, Italy
Ljiljana Zivkovic	Republic Geodetic Authority, Serbia
Luigi Santopietro	Institute of Methodologies for Environmental Analysis National Research Council of Italy CNR IMAA, Italy
Ilaria Delponte	University of Genoa, Italy
Carmen Guida	Università degli Studi di Napoli Federico II, Italy
Chiara Di Dato	University of L'Aquila, Italy

Numerical Aspects in Machine Learning For Dynamical Systems (NAMLYS 2026)

Workshop Organizers

Pasquale De Luca	Parthenope University of Naples, Italy
Livia Marcellino	University of Naples Parthenope, Italy
Valeria Mele	University of Naples Federico II, Italy

Workshop Program Committee Members

Dario Caramiello	University of Naples Parthenope, Italy
Giuseppe Fiorillo	Parthenope University of Naples, Italy
Marco Lettiero	University of Naples Parthenope, Italy
Francesco Peluso	University of Naples Parthenope, Italy

Numerical Methods and Machine Learning in Science & Engineering (NMLSE 2026)

Workshop Organizers

Cecília Coelho	Helmut Schmidt University, Germany
M. Fernanda P. Costa	University of Minho, Portugal

Luís Ferrás	University of Minho, Portugal
Eliana Costa e Silva	Polytechnic of Porto, Portugal

Workshop Program Committee Members

Osvaldo Gramaxo de Freitas	Universitat de València, Spain
Estela Bicho	University of Minho, Portugal
Flora Ferreira	University of Porto, Portugal
Bruno Veloso	Universidade do Porto, Portugal
Luís Louro	University of Minho, Portugal
Weronika Wojtak	CCG/ZGDV Institute, Portugal
Wolfram Erlhagen	University of Minho, Portugal

6th Workshop on Privacy in the Cloud/Edge/IoT World 2026 (PCEIoT 2026)

Workshop Organizers

Lelio Campanile	Università degli Studi della Campania Luigi Vanvitelli, Italy
Mauro Iacono	Università degli Studi della Campania Luigi Vanvitelli, Italy
Michele Mastroianni	University of Foggia, Italy

Workshop Program Committee Members

Fiammetta Marulli	University of Campania Luigi Vanvitelli, Italy
Christian Riccio	Università degli Studi della Campania Luigi Vanvitelli, Italy
Arcangelo Castiglione	University of Salerno, Italy
Gianni D'Angelo	University of Salerno, Italy
Biagio Boi	University of Salerno, Italy
Luigi Piero Di Bonito	Università degli Studi della Campania, Italy

Preserving Our Past: Spatial and Remote Sensing Technologies for Cultural Heritage in a Changing Climate (POP 2026)

Workshop Organizers

Maria Danese	Institute of Heritage Science of the National Research Council of Italy (CNR ISPC), Italy
Rosa Lasaponara	Institute of Methodologies for Environmental Analysis (CNR IMAA), Italy
Nicola Masini	Institute of Heritage Science of the National Research Council of Italy (CNR ISPC), Italy

Workshop Program Committee Members

Dario Gioia	Institute of Heritage Science of the National Research Council of Italy (CNR ISPC), Italy
Marilisa Biscione	Institute of Heritage Science of the National Research Council of Italy (CNR ISPC), Italy
Valentina Florio	CNR, Italy
Antonio Minervino Amodio	Institute of Heritage Science of the National Research Council of Italy (CNR ISPC), Italy

Scientific Computing Infrastructure (SCI 2026)

Workshop Organizers

Vladimir Korkhov	Saint Petersburg State University, Russia
Elena Stankova	Saint Petersburg State University, Russia

Workshop Program Committee Members

Adam Belloum	University of Amsterdam, The Netherlands
Dmitrii Vasiunin	Deutsche Telekom Cloud Services E.P.E., Greece
Serob Balyan	Institute for Informatics and Automation Problems, National Academy of Sciences of Armenia, Armenia
Suren Abrahamyan	Institute for Informatics and Automation Problems of NAS RA, Armenia
Ashot Gevorkyan	Institute for Informatics and Automation Problems NAS of Armenia, Armenia

Michal Hnatič	P.J. Šafárik University in Košice, Slovak Republic
Martin Vala	Pavol Jozef Šafárik University, Slovak Republic
Nodir Zaynalov	Samarkand Branch of Tashkent University of Information Technologies Named After Muhammad al Khwarizmi, Uzbekistan
Michail Panteleyev	Saint Petersburg Electrotechnical University, Russia
Alexander Degtyarev	Saint Petersburg State University, Russia
Alexander Bogdanov	St. Petersburg State University, Russia
Nataliia Kulabukhova	AutoTech, Russia
Alexander Degtyarev	St.Petersburg State University, Russia
Alexander Bogdanov	St. Petersburg State University, Russia

Socio-Economic and Environmental Models for Land Use Management (SEMLUM 2026)

Workshop Organizers

Debora Anelli	Sapienza University of Rome, Italy
Pierluigi Morano	Polytechnic University of Bari, Italy
Francesco Tajani	Sapienza University of Rome, Italy
Felicia Di Liddo	Polytechnic University of Bari, Italy

Workshop Program Committee Members

Benedetto Manganelli	Università degli Studi della Basilicata, Italy
Emma Sabatelli	Sapienza Università di Roma, Italy
María Paz Sáez-Pérez	University of Granada, Spain
Yasmine Selim	Ain Shams University, Egypt
Daniela Tavano	University of Calabria, Italy
Francesco Calabrò	Mediterranea University of Reggio Calabria, Italy
Pierfrancesco De Paola	University of Naples Federico II, Italy
Francesco Paolo Del Giudice	University of Rome La Sapienza, Italy
Paola Amoruso	LUM Giuseppe Degennaro University, Italy
Rubina Canesi	Università degli Studi di Padova, Italy
Daniele La Rosa	University of Catania, Italy
Giuseppe Cerullo	Sapienza University of Rome, Italy
Antonella Roma	Sapienza University of Rome, Italy
Francesca Fariello	Sapienza University of Rome, Italy

Ports and logistics of the future - Smartness and Sustainability (SmartPorts 2026)

Workshop Organizers

Andrea Gallo	Università degli Studi di Trieste, Italy
Giuseppe Borruso	University of Trieste, Italy
Gianfranco Fancello	University of Cagliari, Italy
Veronica Camerada	Università degli studi di Sassari, Italy
Clara Di Fazio	University of Trieste, Italy

Workshop Program Committee Members

Patrizia Serra	University of Cagliari, Italy
Francesco Bruzzone	Università Iuav di Venezia, Italy
Marco Mazzarino	University Iuav of Venice, Italy
Maria del Mar Munoz	University of Cádiz, Spain
Ainhoa Amaro	EVOCONS, Spain
Brunella Brundu	University of Sassari, Italy
Giovanni Mauro	University of Campania Luigi Vanvitelli, Italy
Maria Ronza	University of Naples Federico II, Italy
Marcello Tadini	Università del Piemonte Orientale, Italy
Luca Toneatti	University of Trieste, Italy
Mara Ladu	University of Cagliari, Italy
Ginevra Balletto	University of Cagliari, Italy
Francesca Sinatra	University of Trieste, Italy

Sustainable evolution of long-Distance frEight and paSsenger Transport (SOLIDEST 2026)

Workshop Organizers

Antonio Russo	University of Enna Kore, Italy
Giuseppe Musolino	Università Mediterranea di Reggio Calabria, Italy
Socrates Basbas	Aristotle University of Thessaloniki, Greece
Oana Dinu	National University of Science and Technology Politehnica Bucharest, Romania
Sanja Šurdonja	University of Rijeka, Croatia
Elen Twrdy	University of Ljubljana, Slovenia

Workshop Program Committee Members

Tiziana Campisi	University of Enna Kore, Italy
Massimo Di Gangi	Università di Messina, Italy
Antonio Polimeni	Università di Messina, Italy
Efstathios Bouhouras	Aristotle University of Thessaloniki, Greece
Florin Rusca	National University of Science and Technology Politehnica Bucharest, Romania
Vittorio Astarita	University of Calabria, Italy
Giuseppe Guido	Università della Calabria, Italy
Stefanos Tsigdinos	National Technical University of Athens, Greece
Chiara Garau	University of Cagliari, Italy

Sustainability Performance Assessment: Models, Approaches, and Applications Toward Interdisciplinary and Integrated Solutions (SPA 2026)

Workshop Organizers

Francesco Scorza	University of Basilicata, Italy
Sabrina Lai	University of Cagliari, Italy
Jolanta Dvarioniene	Kaunas University of Technology, Lithuania
Michele Campagna	University of Cagliari, Italy

Workshop Program Committee Members

Corrado Zoppi	University of Cagliari, Italy
Maria Cerreta	University of Naples Federico II, Italy
Giuseppe Borruso	University of Trieste, Italy
Chiara Garau	University of Cagliari, Italy
Ginevra Balletto	University of Cagliari, Italy
Alenka Poplin	Iowa State University, USA
Francesco Zullo	University of L'Aquila, Italy
Alessandro Marucci	University of L'Aquila, Italy
Federica Leone	University of Cagliari, Italy
Simone Corrado	University of Basilicata, Italy
Alfonso Annunziata	University of Basilicata, Italy
Rachele Vanessa Gatto	Università della Basilicata, Italy
Francesco Paolo Del Giudice	University of Basilicata, Italy

Specifics of smart cities development in Europe (SPEED 2026)

Workshop Organizers

Chiara Garau	University of Cagliari, Italy
Katarína Vitálišová	Matej Bel University, Slovak Republic
Marco Fanfani	University of Florence, Italy
Anna Vaňová	Matej Bel University, Slovak Republic
Kamila Borsekova	Matej Bel University, Slovak Republic
Paola Zamperlin	University of Pisa, Italy

Workshop Program Committee Members

Francesca Maltinti	University of Cagliari, Italy
Silvia Rossetti	Università degli Studi di Parma, Italy
Carmen Guida	Università degli Studi di Napoli Federico II, Italy
Floriana Zucaro	University of Naples Federico II, Italy
Giovanna Concu	University of Cagliari, Italy
Paolo Nesi	University of Florence, Italy
Pierfrancesco Bellini	University of Florence, Italy
Gerardo Carpentieri	University of Naples Federico II, Italy
Radoslav Kožiak	Matej Bel University, Slovak Republic
Lucia Dobrotová	Matej Bel University in Banská Bystrica, Slovak Republic
Veronika Piscová	Institute of Landscape Ecology Slovak Academy of Sciences, Slovak Republic
Darina Rojíková	Matej Bel University, Slovak Republic
Katarína Sýkorová	Matej Bel University, Slovak Republic
Silvia Ondrejičková	Comenius University Bratislava, Slovak Republic
Vladimír Ondrejička	Comenius University Bratislava, Slovak Republic
Maros Finka	Comenius University Bratislava, Slovak Republic
Jan Nevima	Masaryk University, Czech Republic
Christian Corral Burau	Leibniz University Hannover, Germany
Kamila Turečková	Silesian University in Opava, Czech Republic
Luciano Alessandro Ipsaro Palesi	University of Florence, Italy
Enrico Collini	University of Florence, Italy
Gianni Pantaleo	University of Florence, Italy
Krishnamoorthy Ranganathan	Chennai Institute of Technology, India

Smart, Safe, and Health Cities (SSHC 2026)

Workshop Organizers

Chiara Garau	University of Cagliari, Italy
Gerardo Carpentieri	University of Naples Federico II, Italy
Carmen Guida	Università degli Studi di Napoli Federico II, Italy
Martina Carra	University of Brescia, Italy
Vincenza Torrisi	University of Catania, Italy
Silvia Rossetti	Università degli Studi di Parma, Italy

Workshop Program Committee Members

Romano Fistola	University of Naples Federico II, Italy
Francesca Maltinti	University of Cagliari, Italy
Barbara Caselli	Università degli Studi di Parma, Italy
Tiziana Campisi	University of Enna Kore, Italy
Katarína Sýkorová	Matej Bel University, Slovak Republic
Tazyeen Alam	Università degli Studi di Cagliari, Italy
Francesco Zullo	University of L'Aquila, Italy
Giovanna Concu	University of Cagliari, Italy
Anastasia Stratigea	National Technical University of Athens, Greece
Ezio Riggi	National Research Council, Italy
Miriam Distefano	CNR, Italy

Smart and Sustainable Island Communities (SSIC 2026)

Workshop Organizers

Chiara Garau	University of Cagliari, Italy
Anastasia Stratigea	National Technical University of Athens, Greece
Yiota Theodora	National Technical University of Athens, Greece
Giovanna Concu	University of Cagliari, Italy
Tiziana Campisi	University of Enna Kore, Italy
Tanja Congiu	University of Sassari, Italy

Workshop Program Committee Members

Francesca Maltinti	University of Cagliari, Italy
Mauro Coni	University of Cagliari, Italy
Paola Zamperlin	University of Florence, Italy

Tarek Teba	University of Portsmouth, UK
Gerardo Carpentieri	University of Naples Federico II, Italy
Floriana Zucaro	University of Naples Federico II, Italy
Roberto Minunno	Curtin University, Australia
Apostolos Lagarias	University of Thessaly, Greece
Giuseppina Vacca	University of Cagliari, Italy
Carmen Guida	Università degli Studi di Napoli Federico II, Italy
Tazyeen Alam	Università degli Studi di Cagliari, Italy
Mauro D'Apuzzo	University of Cassino and Southern Lazio, Italy
Mana Dastoum	Polytechnic University of Madrid, Spain
Alfonso Annunziata	University of Basilicata, Italy
Barbara Caselli	Università degli Studi di Parma, Italy
Silvia Rossetti	Università degli Studi di Parma, Italy
Gian Felice Giaccu	University of Sassari, Italy
Anastasia Stratigea	National Technical University of Athens, Greece
Vincenza Torrisi	University of Catania, Italy

Sustainable Tourism Economy, Silk Road, Reuse and Enhancement of Cultural Heritage (STESilkRoad 2026)

Workshop Organizers

Celestina Fazia	Università di Enna Kore, Italy
Alexandru Ilies	University of Oradea, Romania
Dorina Camelia Ilies	University of Oradea, Romania
Cristina Natoli	Ministero dell'Ambiente e della Sicurezza Energetica, Italy
Bahodirhon Safarov	Samarkand Branch of Tashkent State University of Economics, Uzbekistan
Clara Stella Vicari Aversa	Università degli Studi Mediterranea di Reggio Calabria, Italy

Workshop Program Committee Members

Giulia Fernanda Grazia Catania	Università di Enna Kore, Italy
Federica Sortino	Università degli Studi di Enna Kore, Italy
Thowayeb Hassan	King Faisal University, Saudi Arabia
Tatar Corina	University of Oradea, Romania
Maria Francesca Faro	Ministry of Education and Merit (MIM), Italy

Theoretical and Computational Chemistry and its Applications (TCCM 2026)

Workshop Organizers

Noelia Faginas Lago	Università di Perugia, Italy
Andrea Lombardi	Università di Perugia, Italy
Marcos Mandado Alonso	University of Vigo, Spain

Workshop Program Committee Members

Jose Manuel Hermida Ramon	Universidade de Vigo, Spain
Luca Mancini	University of Perugia, Italy
Stefano Falcinelli	University of Perugia, Italy
Ángel Martín Pendás	Universidad de Oviedo, Spain
Enrique M. Cabaleiro Lago	Universidade de Santiago de Compostela, Spain
Giacomo Giorgi	Università degli Studi di Perugia, Italy
Stefano Evangelisti	Université de Toulouse, France
Nicolás Ramos Berdullas	University of Vigo, Spain
Juan José Nogueira	Universidad Autónoma de Madrid, Spain
Giacomo Pannacci	Università degli Studi di Perugia, Italy

Transport Infrastructures for Smart Cities (TISC 2026)

Workshop Organizers

Francesca Maltinti	University of Cagliari, Italy
Mauro Coni	University of Cagliari, Italy
Chiara Garau	University of Cagliari, Italy
Benedetto Barabino	University of Brescia, Italy
Nicoletta Rassu	Università di Cagliari, Italy
James Rombi	University of Cagliari, Italy

Workshop Program Committee Members

Tiziana Campisi	University of Enna Kore, Italy
Mauro D'Apuzzo	University of Cassino and Southern Lazio, Italy
Roberto Ventura	Università degli Studi di Brescia, Italy
Roberto Minunno	Curtin University, Australia

Giovanna Concu	University of Cagliari, Italy
Claudia Loggia	University of KwaZulu-Natal, South Africa

Transforming Urban Analytics: The Impact of Crowdsourced Mapping and Advanced AI Techniques on Future Cities (Tr-UrbAna 2026)

Workshop Organizers

Ayse Giz Gulnerman Gengec	Ankara Hacı Bayram Veli University, Turkey
Müslüm Hacar	Yildiz Technical University, Turkey

Workshop Program Committee Members

Beniamino Murgante	University of Basilicata, Italy
Abdulkadir Memduhoğlu	Harran University, Turkey
Zeynel Abidin Polat	Izmir Katip Celebi University, Turkey
Güzide Miray Perihanoglu	Van Yuzuncu Yil University, Turkey
Tugba Memisoglu Baykal	Hacı Bayram Veli University, Turkey
Federico Mara	University of Pisa, Italy
Himmet Karaman	Istanbul Technical University, Turkey
Özge Öztürk Hacar	Yildiz Technical University, Turkey

From structural to TRAnsformative-change of City Environment: challenges & solutions & perspectives (TRACE 2026)

Workshop Organizers

Maria Rosaria Guarini	Sapienza University of Rome, Italy
Pierluigi Morano	Polytechnic University of Bari, Italy
Francesco Sica	Sapienza University of Rome, Italy
Alejandro Segura	Universidad Politécnica de Madrid, Spain
Yilsy María Nuñez Guerrero	Universidad Politécnica de Madrid, Spain

Workshop Program Committee Members

Maria Cerreta	University of Naples Federico II, Italy
Francesco Scorza	University of Basilicata, Italy

Francesco Calabrò	Mediterranea University of Reggio Calabria, Italy
Felicia Di Liddo	Polytechnic University of Bari, Italy
Marco Locurcio	Polytechnic University of Bari, Italy
Valeria Saiu	University of Cagliari, Italy
Lucia Della Spina	University Mediterranea of Reggio Calabria, Italy
Fabrizio Battisti	University of Florence, Italy
Giuseppe Cerullo	Sapienza University of Rome, Italy
Fabiana Forte	Università della Campania Luigi Vanvitelli, Italy
Marta Dell'Ovo	Politecnico di Milano, Italy
Cristina Coscia	Politecnico di Torino, Italy
Debora Anelli	Sapienza University of Rome, Italy

Urban Artificial Intelligence for Modelling, Analysis & Planning (Urban plAnnIng 2026)

Workshop Organizers

Chiara Garau	University of Cagliari, Italy
Romano Fistola	University of Naples Federico II, Italy
Beniamino Murgante	University of Basilicata, Italy
Dessislava Petrova Antonova	Sofia University, GATE Institute, Bulgaria
Gianni Pantaleo	University of Florence, Italy
Rosa Anna La Rocca	University of Naples Federico II

Workshop Program Committee Members

Paola Zamperlin	University of Florence, Italy
Francesco Zullo	University of L'Aquila, Italy
Gerardo Carpentieri	University of Naples Federico II, Italy
Luciano Alessandro Ipsaro Palesi	University of Florence, Italy
Tazyeen Alam	Università degli Studi di Cagliari, Italy
Reza Askarizad	University of Cagliari, Italy
Mana Dastoum	Polytechnic University of Madrid, Spain
Alfonso Annunziata	University of Basilicata, Italy

Urban Space Accessibility and Mobilities (USAM 2026)

Workshop Organizers

Chiara Garau	University of Cagliari, Italy
Alessandro Plaisant	University of Sassari, Italy
Barbara Caselli	Università degli Studi di Parma, Italy
Mauro D'Apuzzo	University of Cassino and Southern Lazio, Italy
Francesco Pinna	University of Cagliari, Italy
Silvio Nocera	IUAV University of Venice, Italy

Workshop Program Committee Members

Mauro Coni	University of Cagliari, Italy
Martina Carra	University of Brescia, Italy
Tiziana Campisi	University of Enna Kore, Italy
Reza Askarizad	University of Cagliari, Italy
Giuseppe Cappelli	University of Cassino and Southern Lazio, Italy
Floriana Zucaro	University of Naples Federico II, Italy
Gerardo Carpentieri	University of Naples Federico II, Italy
Youssef El Ganadi	International University of Rabat, Morocco
Yiota Theodora	National Technical Univeristy of Athens, Greece
Silvia Rossetti	Università degli Studi di Parma, Italy
Aime Tsinda	University of Rwanda, Rwanda
Paola Zamperlin	University of Florence, Italy
Luciano Alessandro Ipsaro Palesi	University of Florence, Italy
Tazyeen Alam	Università degli Studi di Cagliari, Italy
Mariano Pernetti	Università della Campania Luigi Vanvitelli, Italy
Salvatore Leonardi	University of Catania, Italy
Francesco Galante	University of Naples Federico II, Italy

Virtual Reality, Augmented Reality and Applications (VRA 2026)

Workshop Organizers

Damiano Perri	University of Perugia, Italy
Osvaldo Gervasi	University of Perugia, Italy
Paolo Nesi	University of Florence, Italy

Workshop Program Committee Members

Taihoon Kim	Chonnam National University, South Korea
JungYoon Kim	Gachon University, South Korea
David Berti	ART SpA, Italy

Workshop on Advanced and Computational Methods for Earth Science applications (WACM4ES 2026)

Workshop Organizers

Luca Piroddi	University of Cagliari, Italy
Patrizia Caprizzi	University of Palermo, Italy
Marilena Cozzolino	Università del Molise, Italy
Giuseppina Vacca	University of Cagliari, Italy
Luigi Capozzoli	Hydrogeosite Laboratory of CNR-IMAA, Italy
Nikos Papadopoulos	Foundation for Research and Technology Hellas, Greece

Workshop Program Committee Members

Sebastiano D'Amico	University of Malta, Malta
Enrica Vecchi	University of Cagliari, Italy
Donatella Rita Fiorino	Università degli Studi di Cagliari, Italy
Chiara Garau	University of Cagliari, Italy
Emanuele Colica	University of Malta, Malta
Raffaele Martorana	Università di Palermo, Italy
Raffaele Persico	Università della Calabria, Italy
Andrea Angelini	Institute of Heritage Science of the National Research Council of Italy (CNR ISPC), Italy
Paolo Mauriello	Università del Molise, Italy
Dedalo Marchetti	Italian National Institute of Geophysics and Volcanology, Italy
Veronica Pazzi	University of Florence, Italy
Agnese Innocenti	University of Florence, Italy
Maurizio Ercoli	University of Perugia, Italy
Luciano Galone	University of Malta, Malta
Amr Abd El-Raouf	Zagazig University, Egypt
Peter Iregbeyen	University of Malta, Malta
Pietro Milillo	University of Houston, USA
Tommaso Beni	University of Florence, Italy

Cristiano Fidani	Central Italy Electromagnetic Network, Italy
Deodato Tapete	Italian Space Agency (ASI), Italy
Enzo Rizzo	University of Ferrara, Italy
Guido Luzi	Centre Tecnològic de Telecomunicacions de Catalunya, Spain
Sina Saneiyan	Binghamton University, USA
Alex Sendrós	Universitat de Barcelona, Spain
Mercedes Solla	Universidade de Vigo, Spain
Alessandra Carollo	University of Palermo, Italy
Andrea Vergnano	Università di Torino, Italy
Giorgio Alaia	Università di Perugia, Italy
Giorgio De Donno	Sapienza University of Rome, Italy
Samuel Bignardi	University of Chieti and Pescara, Italy
Athos Agapiou	Cyprus University of Technology, Cyprus
Gregory Tsokas	Aristotle University of Thessaloniki, Greece
Alexandru Hegyi	Oslo University, Norway
Ilaria Barone	Università degli Studi di Padova, Italy
Yuan Qi	University of Hong Kong, China
Carmen Cuenca García	Universitat de València, Spain
Ilaria Catapano	Institute for Electromagnetic Sensing of the Environment (IREA) CNR, Italy
Mariano Lisi	University of Basilicata, Italy
Adrien Dimech	École de Technologie Supérieure, Canada
Davide Campo	Independent Researcher, UK
Chiara Torre	University of Malta, Malta
Anna Forte	University of Bologna, Italy
Adam Gauci	University of Malta, Malta
Vittoria Borghese	TNO, The Netherlands
Andrea Scianna	Institute for High Performance Computing and Networking (ICAR) CNR, Italy
Maria Antonia Brovelli	Politecnico di Milano, Italy
Pietro Milillo	University of Houston, USA
Gregory Tsokas	Aristotle University of Thessaloniki, Greece

Sponsoring Organizations

ICCSA 2026 would not have been possible without tremendous support from many organizations and institutions, for which all organizers and participants of ICCSA 2026 express their sincere gratitude:

University of Minho, Portugal (https://www.uminho.pt/)

University of Perugia, Italy (https://www.unipg.it)

University of Basilicata, Italy (http://www.unibas.it)

University of Cagliari, Italy (https://unica.it/)

Monash University, Australia (https://www.monash.edu/)

Kyushu Sangyo University, Japan (https://www.kyusan-u.ac.jp/)

Venue

ICCSA 2026 took place at: **the University of Minho campus:**
Campus de Gualtar
4710-057 Braga
Portugal

Additional Reviewers

Amr Abd Elraouf	Zagazig University, Egypt
Lidia Aceto	Università del Piemonte Orientale, Italy
Ana Maria Acu	Lucian Blaga University of Sibiu, Romania
Vera Afreixo	University of Aveiro, Portugal
Akif Akgul	Hitit University, Turkey
Sania Akhtar	Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Pakistan
Giorgio Alaia	Università di Perugia, Italy
Tazyeen Alam	Università degli Studi di Cagliari, Italy
Arnulfo Alanis	Tecnológico Nacional de México, Mexico
Sandra Aleixo	University of Lisbon, Portugal
Earl Ryan Aleluya	Mindanao State University Iligan Institute of Technology, Philippines
Muhammad Junaid Ali	Université de Haute Alsace, France
Filipe Alvelos	Universidade do Minho, Portugal
Claudia Amadei	University of Padua, Italy
Paulo Amorim	Escola de Matemática Aplicada da Fundação Getulio Vargas, Brazil
Ana Paula Amorim	Universidade do Minho, Portugal
Paola Amoruso	LUM Giuseppe Degennaro University, Italy
Debora Anelli	Sapienza University of Rome, Italy
Mariarosaria Angrisano	Pegaso University, Italy
Alfonso Annunziata	University of Basilicata, Italy
Chiara Anselmi	Università di Pisa, Italy
Shigeyoshi Aoyama	KEK, Japan
Tatsumi Aoyama	University of Tokyo, Japan
Bernady O. Apduhan	Kyushu Sangyo University, Japan
Pablo Luiz Araujo Munhoz	Federal University of Ouro Preto, Brazil
Muhammad Waqas Arshad	University of Bologna, Italy
Daniela Ascenzi	Università degli Studi di Trento, Italy
Laura Ascione	Università degli Studi di Napoli Federico II, Italy
Reza Askarizad	University of Cagliari, Italy
Vanessa Assumma	University of Bologna, Italy
Vittorio Astarita	University of Calabria, Italy
Maria Fernanda Augusto	BITrum Research Center, Spain
Asja Aulisio	Politecnico di Torino, Italy
Samat Baigereyev	D. Serikbayev East Kazakhstan Technical University, Kazakhstan
Marco Baioletti	University of Perugia, Italy
Sara Balderas Díaz	University of Cádiz, Spain

Ginevra Balletto	University of Cagliari, Italy
Serob Balyan	Institute for Informatics and Automation Problems, National Academy of Sciences of Armenia, Armenia
Benedetto Barabino	University of Brescia, Italy
Sebastiano Barbieri	Politecnico di Torino, Italy
Ilaria Barone	Università degli Studi di Padova, Italy
Alice Barreca	Politecnico di Torino, Italy
Socrates Basbas	Aristotle University of Thessaloniki, Greece
Rosaria Battarra	Consiglio Nazionale delle Ricerche, Italy
Silvia Battino	University of Sassari, Italy
Fabrizio Battisti	University of Florence, Italy
Patricia Beites	University of Beira Interior, Portugal
David Berti	Art Spa, Italy
Samuel Bignardi	University of Chieti and Pescara, Italy
Francesco Bimbo	University of Foggia, Italy
Francesca Biondi	Politecnico di Torino, Italy
Marilisa Biscione	CNR, Italy
Alexander Bogdanov	St. Petersburg State University, Russia
Biagio Boi	University of Salerno, Italy
Alessandro Bombini	Istituto Nazionale di Fisica Nucleare, Italy
Ana Borges	Instituto Politécnico do Porto, Portugal
Vittoria Borghese	TNO, The Netherlands
Giuseppe Borruso	University of Trieste, Italy
Kamila Borsekova	Matej Bel University, Slovak Republic
Efstathios Bouhouras	Aristotle University of Thessaloniki, Greece
Ana Cristina Braga	University of Minho, Portugal
Diego Brandão	Federal Center for Technological Education of Rio de Janeiro, Brazil
Maria Antonia Brovelli	Politecnico di Milano, Italy
Francesco Bruzzone	Università Iuav di Venezia, Italy
Francesca Buglione	CNR ISPC, Italy
Enrique M. Cabaleiro Lago	Universidade de Santiago de Compostela, Spain
Francesco Calabrò	Mediterranea University of Reggio Calabria, Italy
Rogerio Calazan	Instituto de Estudos do Mar Almirante Paulo Moreira - Brazilian Navy, Brazil
Lelio Campanile	Università degli Studi della Campania Luigi Vanvitelli, Italy
Tiziana Campisi	University of Enna Kore, Italy
Dario Campisi	University of Perugia, Italy
Rubina Canesi	Università degli Studi di Padova, Italy
Alessandra Capolupo	Polytechnic University of Bari, Italy

Caterina Caprioli	Politecnico di Torino, Italy
Dario Caramiello	University of Naples Parthenope, Italy
Gerardo Carpentieri	University of Naples Federico II, Italy
Martina Carra	University of Brescia, Italy
Barbara Caselli	Università degli Studi di Parma, Italy
Danny Casprini	Politecnico di Milano, Italy
Arcangelo Castiglione	University of Salerno, Italy
Giulia Fernanda Grazia Catania	Università di Enna Kore, Italy
Ilaria Catapano	CNR IREA, Italy
Laura Cavalcante	Instituto Superior de Engenharia do Porto, Portugal
Giulio Cavana	Politecnico di Torino, Italy
Valeria Cera	University of Naples Federico II, Italy
Maria Cerreta	University of Naples Federico II, Italy
Giuseppe Cerullo	Sapienza University of Rome, Italy
Daniele Cesini	CNAF, INFN, Italy
Kai Cheng	Kyushu Sangyo University, Japan
Giuseppe Ciciriello	University Of Naples Federico II, Italy
Birol Ciloglugil	Ege University, Turkey
Cecília Coelho	Helmut Schmidt University, Germany
Enrico Collini	University of Florence, Italy
Luís Conceicao	Instituto Superior de Engenharia do Porto, Portugal
Giovanna Concu	University of Cagliari, Italy
Tatar Corina	University of Oradea, Romania
Alejandro Cornejo	Universitat Politècnica de Catalunya, Spain
Simone Corrado	University of Basilicata, Italy
Florbela Correia	Polytechnic Institute of Viana do Castelo, Portugal
Franco Corti	University of Padua, Italy
Cristina Coscia	Politecnico di Torino, Italy
Lino Costa	University of Minho, Portugal
Eliana da Costa e Silva	Polytechnic Institute of Porto, Portugal
Alessandro Costantini	INFN, Italy
Alessia Crisopulli	Politecnico di Milano, Italy
Valerio Cutini	University of Pisa, Italy
Mauro D'Apuzzo	University of Cassino and Southern Lazio, Italy
Hiroshi Daisaka	Hitotsubashi University, Japan
Sebastiano Damico	University of Malta, Malta
Maria Danese	CNR ISPC, Italy
Gianni Dangelo	University of Salerno, Italy
Mana Dastoum	Polytechnic University of Madrid, Spain

Giulia Datola	Politecnico di Milano, Italy
Regina De Almeida	University of Trás-os-Montes and Alto Douro, Portugal
Elena De Angelis	Politecnico di Torino, Italy
Elise de Doncker	Western Michigan University, USA
Pasquale De Luca	Parthenope University of Naples, Italy
Luiza De Macedo Mourelle	State University of Rio de Janeiro, Brazil
Pierfrancesco De Paola	University of Naples Federico II, Italy
Alexander Degtyarev	St.Petersburg State University, Russia
Francesco Paolo Del Giudice	University of Basilicata, Italy
Piera Della Morte	Pegaso Telematic University, Italy
Federico Dellanna	Politecnico di Torino, Italy
Marta Dellovo	Politecnico di Milano, Italy
Federica Deri	Università di Pisa, Italy
Francesco Di Capua	University of Basilicata, Italy
Clara Di Fazio	University of Trieste, Italy
Gabriella Di Genova	Università degli Studi di Perugia, Italy
Benedetta Di Leo	Università Politecnica delle Marche, Italy
Felicia Di Liddo	Polytechnic University of Bari, Italy
Joana Dias	University of Coimbra, Portugal
Oana Dinu	National University of Science and Technology Politehnica Bucharest, Romania
Lucia Dobrotová	Matej Bel University in Banská Bystrica, Slovak Republic
Edda Donati	University of Rome Tor Vergata, Italy
Aziz Dursun	Virginia Tech University, USA
Zahra Ebadi Ansaroudi	Fondazione Bruno Kessler, Italy
Maurizio Ercoli	University of Perugia, Italy
Dario Esposito	Polytechnic University of Bari, Italy
Noelia Faginas Lago	Università di Perugia, Italy
Maria Irene Falcao	University of Minho, Portugal
Stefano Falcinelli	University of Perugia, Italy
Marco Fanfani	University of Florence, Italy
Marcos Farias	Instituto de Engenharia Nuclear, Brazil
Francesca Fariello	Sapienza University of Rome, Italy
Maria Francesca Faro	Ministry of Education and Merit (MIM), Italy
Eslam Farsimadan	University of Salerno, Italy
Sara Favargiotti	University of Trento, Italy
Celestina Fazia	Università di Enna Kore, Italy
Maksims Feofilovs	Riga Technical University, Latvia
Florbela P. Fernandes Cedri	Instituto Politecnico de Bragança, Portugal
Paula Odete Fernandes	Instituto Politecnico de Bragança, Portugal

Flora Ferreira	University of Porto, Portugal
Maddalena Ferretti	Università Politecnica delle Marche, Italy
Luís Ferrás	University of Porto, Portugal
Massimo Ficco	University of Salerno, Italy
Maros Finka	Masaryk University, Slovak Republic
Ugo Fiore	University of Salerno, Italy
Lorena Fiorini	University of L'Aquila, Italy
Romano Fistola	University of Naples Federico II, Italy
Beatriz Flamia Azevedo	Instituto Politecnico de Bragança, Portugal
Valentina Florio	CNR, Italy
Ali Forooghi	University of Campania Luigi Vanvitelli, Italy
Anna Forte	University of Bologna, Italy
Grigorios Fountas	Aristotle University of Thessaloniki, Greece
Andreas Fricke	Hasso Plattner Institute for Digital Engineering, University of Potsdam, Germany
Luigi Fusco Girard	Pegaso Telematic University, Italy
Francesco Galante	University of Naples Federico II, Italy
Stefano Galassi	University of Florence, Italy
Andrea Gallo	Università degli Studi di Trieste, Italy
Chiara Garau	University of Cagliari, Italy
Gustavo Gatica	Universidad Andrés Bello, Chile
Rachele Vanessa Gatto	Università della Basilicata, Italy
Ivan Gerace	Università degli Studi di Perugia, Italy
Amal Gerges	University of Cagliari, Italy
Osvaldo Gervasi	University of Perugia, Italy
Tasnim Gharaibeh	Kalamazoo College, USA
Gian Felice Giaccu	University of Sassari, Italy
Dario Gioia	CNR ISPC, Italy
Luca Giommi	INFN CNAF, Italy
Valentina Giorgetti	Università degli Studi di Perugial, Italy
Giacomo Giorgi	Università degli Studi di Perugia, Italy
Hiago Mayk Gomes De Araújo Rocha	National Laboratory for Scientific Computing, Brazil
Omar Gomez	Universidad Nacional de Colombia, Colombia
A. Manuela Goncalves	University of Minho, Portugal
Joaquim Goncalves	Polytechnic Institute of Cávado and Ave, Portugal
Pedro Henrique Gonzalez	Federal University of Rio de Janeiro, Brazil
Yusuke Gotoh	Okayama University, Japan
Osvaldo Gramaxo De Freitas	Universitat de València, Spain
Antonia Gravagnuolo	National Research Council Italy (CNR) Institute of Heritage Science, Italy

Teresa Guarda	Universidad Estatal Península De Santa Elena, Ecuador
Maria Rosaria Guarini	Sapienza University of Rome, Italy
Raphael Guerra	Universidade Federal Fluminense, Brazil
Hélia Guerra	University of the Azores, Portugal
Gabriel Guerrero Contreras	University of Cádiz, Spain
Carmen Guida	Università degli Studi di Napoli Federico II, Italy
Eligius Hendrix	Universidad de Málaga, Spain
Jose Manuel Hermida Ramon	Universidade de Vigo, Spain
Antonino Iannuzzo	Universitá degli Studi del Sannio, Italy
Luciano Alessandro Ipsaro Palesi	University of Florence, Italy
Issaku Kanamori	Riken, Japan
Yeliz Karaca	UMass Chan Medical School and Massachusetts Institute of Technology, USA
Himmet Karaman	Istanbul Technical University, Turkey
Tatyana Khachkova	IPGG SB RAS, Russia
Jungyoon Kim	Gachon University, South Korea
Vladimir Korkhov	Saint Petersburg State University, Russia
Iuliia Kozlova	University of Bologna, Italy
Radoslav KožIak	Matej Bel University, Slovak Republic
Ludovica La Rocca	University of Naples Federico II, Italy
Rosa Anna La Rocca	University of Naples Federico II, Italy
Adnane Labbaci	Ibn Tofail University, Morocco
Sabrina Lai	University of Cagliari, Italy
Giuseppe Francesco Cesare Lama	University of Naples Federico Ii, Italy
Isabella Maria Lami	Politecnico di Torino, Italy
Salvatore Leonardi	University of Catania, Italy
Federica Leone	University of Cagliari, Italy
Vita Leonessa	University of Basilicata, Italy
Barbara Lino	Università degli Studi di Palermo, Italy
Vadim Lisitsa	Institute of Petroleum Geology and Geophysics SB RAS, Russia
Marco Locurcio	Polytechnic University of Bari, Italy
Andrea Lombardi	Università di Perugia, Italy
Cristina Lopes	Polytechnic Institute of Porto, Portugal
Isabel Lopes	Instituto Politécnico de Bragança, Portugal, Italy
Rui Pedro Lopes	Instituto Politécnico de Bragança, Portugal
Nicolas Lori	University of the Azores, Portugal
Luís Louro	University of Minho, Portugal
Jader Lugon Junior	Instituto Federal Fluminense, Brazil
Guido Luzi	Centre Tecnològic de Telecomunicacions de Catalunya, Spain

Antonio S. Macias	University of A Coruña, Spain
Daniela Maiullari	Delft University of Technology, The Netherlands
Robert Makin	Western Michigan University, USA
Giorgia Malavasi	Politecnico di Torino, Italy
Francesca Maltinti	University of Cagliari, Italy
Luca Mancini	University of Perugia, Italy
Marcos Mandado Alonso	University of Vigo, Spain
Benedetto Manganelli	Università degli Studi della Basilicata, Italy
Federico Mara	University of Pisa, Italy
Livia Marcellino	University of Naples Parthenope, Italy
Dedalo Marchetti	INGV, Italy
Francesco Marchetti	University of Padua, Italy
Alessandra Marra	University of Salerno, Italy
Barbara Martelli	INFN, Italy
Valerio Martinelli	Università degli Studi di Napoli Federico II, Italy
Ángel Martín Pendás	Universidad de Oviedo, Spain
Alessandra Mascitelli	University of Chieti Pescara, Italy
Michele Mastroianni	University of Foggia, Italy
Luís Miguel Matos	Universidade do Minho, Portugal
Hideo Matsufuru	High Energy Accelerator Research Organization, Japan
Luis Mazon	BITrum Research Group, Spain
Beatrice Mecca	Politecnico di Torino, Italy
Valeria Mele	University of Naples Federico II, Italy
Abdulkadir Memduhoğlu	Harran University, Turkey
Joao Mendes	Instituto Politecnico de Braganca, Portugal
Luis Mendes Gomes	University of the Azores, Portugal
Paolo Mengoni	Hong Kong Baptist University, China
Daniela Menna	Università degli Studi della Campania Luigi Vanvitelli, Italy
Antonio Minervino Amodio	CNR ISPC, Italy
Roberto Minunno	Curtin University, Australia
Fernando Miranda	Universidade do Minho, Portugal
Andrea Montanino	International Centre for Numerical Methods in Engineering, Spain
Andrea Montanino	University of Naples Parthenope, Italy
Ana Monteiro	University of Coimbra, Portugal
Leila Moradi	University of Salerno, Italy
Joao Morais	ITAM, Mexico
Filipe Mota Pinto	Instituto Politécnico de Leiria, Portugal
Ricardo Moura	NOVA University of Lisbon, Portugal

Maria Mourao	Polytechnic Institute of Viana do Castelo, Portugal
Eugenio Muccio	University of Naples Federico II, Italy
Beniamino Murgante	University of Basilicata, Italy
Giuseppe Musolino	Università Mediterranea di Reggio Calabria, Italy
Naohito Nakasato	University of Aizu, Japan
Isabel Natário	Nova University of Lisbon, Portugal
Nadia Nedjah	State University of Rio de Janeiro, Brazil
Grazia Neglia	Pegaso Telematic University, Italy
Paolo Nesi	University of Florence, Italy
Keigo Nitadori	Riken, Japan
Juan José Nogueira	Universidad Autónoma de Madrid, Spain
Mikhail Novikov	Sobolev Institute of Mathematics SB RAS, Russia
Yilsy María Nuñez Guerrero	Universidad Politécnica de Madrid, Spain
Ana Oliveira	Academia Militar, Portugal
Carlo Olivieri	Pegaso University, Italy
Vladimír Ondrejička	Comenius University in Bratislava, Slovak Republic
Beniamino Onorati	Università degli Studi della Basilicata, Italy
Fabian Bernal Orozco Barrera	UNAM, Mexico
M. Fernanda P. Costa	University of Minho, Portugal
António Pacheco	Instituto Superior Técnico, Portugal
Emerson Paiva	University of Minho, Portugal
Simone Palladino	Università degli Studi della Campania L. Vanvitelli, Italy
Francesco Palmieri	University of Salerno, Italy
Simona Panaro	University of Sussex Business School, UK
Claudia Panciera	IRSA CNR, Italy
Giacomo Pannacci	Università degli Studi di Perugia, Italy
Gianni Pantaleo	University of Florence, Italy
Eric Pardede	La Trobe University, Australia
Rosaria Parente	Universitas Mercatorum, Italy
Vittorio Paris	University of Bergamo, Italy
Luca Parodi	University of Genoa, Italy
Rui Pascoal	University of Coimbra, Portugal
Hugo Peixoto	University of Minho, Portugal
Francesco Peluso	University of Naples Parthenope, Italy
Ana Pereira	Polytechnic Institute of Bragança, Portugal
Amina Pereno	Politecnico di Torino, Italy
Martin Perez Perez	University of Vigo, Spain
Donato Perfetto	University of Campania Luigi Vanvitelli, Italy
Güzide Miray Perihanoglu	Van Yuzuncu Yil University, Turkey

Mariano Pernetti	Università della Campania Luigi Vanvitelli, Italy
Damiano Perri	University of Perugia, Italy
Pasquale Perrini	Consorzio Interuniversitario per l'Idrologia, Italy
Raffaele Persico	Università della Calabria, Italy
Lucia Petronio	University of Bologna, Italy
Rossella Petti	Politecnico di Bari, Italy
Francesco Picariello	Universitas Mercatorum, Italy
Angela Pilogallo	IMAA CNR, Italy
Natalia Pingaro	Politecnico di Milano, Italy
Telmo Miguel Pires Pinto	University of Minho, Portugal
Veronika Piscová	Institute of Landscape Ecology, Slovak Academy of Sciences, Slovak Republic
Zeynel Abidin Polat	Izmir Katip Celebi University, Turkey
Giuliano Poli	Università di Napoli Federico II, Italy
Antonio Polimeni	Università di Messina, Italy
Maurizio Pollino	ENEA, Italy
Mohammad Pourfouladi	Politecnico di Milano, Italy
Dmitry Prokhorov	IPGG SB RAS, Russia
Faiza Qamar	National University of Sciences and Technology, Pakistan
Yuan Qi	University of Hong Kong, China
Siyu Qiu	University of Bologna, Italy
Arunasalam Rahunanthan	Central State University, USA
Joelmir Ramos	Universidade do Estado do Rio de Janeiro, Brazil
Bruna Ramos	University of Minho, Portugal
Nicolás Ramos Berdullas	University of Vigo, Spain
Krishnamoorthy Ranganathan	Chennai Institute of Technology, India
Nicoletta Rassu	Università di Cagliari, Italy
Roberta Ravanelli	University of Liège, Belgium
Michela Ravanelli	Sapienza Università di Roma, Italy
Manuela Rebaudengo	Politecnico di Torino, Italy
Simone Rebegoldi	University of Modena and Reggio Emilia, Italy
Pier Francesco Recchi	Università di Napoli Federico II, Italy
Rommel Regis	Saint Joseph's University, USA
Marco Reis	University of Coimbra, Portugal
James Rhodes	Western Michigan University, USA
Luigi Maciel Ribeiro	State University of Rio de Janeiro, Brazil
Corrado Rindone	Università degli Studi Mediterranea di Reggio Calabria, Italy
Ana Maria A. C. Rocha	Universidade do Minho, Portugal
Humberto Rocha	University of Coimbra, Portugal
Carlos Rodrigues	Instituto Politécnico de Bragança, Portugal

Ivan Rodriguez Conde	University of Arkansas at Little Rock, USA
Darina Rojíková	Matej Bel University, Slovak Republic
Diana Rolando	Politecnico di Torino, Italy
Antonella Roma	Sapienza University of Rome, Italy
James Rombi	University of Cagliari, Italy
Elisabetta Ronchieri	INFN CNAF, Italy
Marzio Rosi	University of Perugia, Italy
Silvia Rossetti	Università degli Studi di Parma, Italy
Marco Rossitti	Politecnico di Milano, Italy
Florin Rusca	National University of Science and Technology Politehnica Bucharest, Romania
Yvonne Russo	University of Campania L. Vanvitelli, Italy
Antonio Russo	University of Enna Kore, Italy
Ana Sofia Rézio	Polytechnic Institute of Lisbon, Portugal
Emma Sabatelli	Sapienza Università di Roma, Italy
Matteo Sacchet	University of Turin, Italy
Sabrina Sacco	University of Naples Federico II, Italy
Lucia Saganeiti	Institute of Methodologies for Environmental Analysis, National Research Council, Italy
Valeria Saiu	University of Cagliari, Italy
Francesco Santini	University of Perugia, Italy
Rafael Santos	National Institute for Space Research (INPE), Brazil
Edcarllos Santos	Universidade Federal do Sul da Bahia, Brazil
Valentino Santucci	Università per Stranieri di Perugia, Italy
Andrea Sarcina	Università di Padova, Italy
Ilaria Scarcelli	Politecnico di Bari, Italy
Andrea Scianna	National Research Council of Italy, Italy
Francesco Scorza	University of Basilicata, Italy
Ester Scotto Di Perta	University of Naples Federico II, Italy
Alejandro Segura	Universidad Politécnica de Madrid, Spain
Yasmine Selim	Ain Shams University, Egypt
Ines Sena Cedri	Instituto Politécnico de Bragança and University of Minho, Portugal
Alex Sendrós	Universitat de Barcelona, Spain
Ricardo Severino	University of Minho, Portugal
Mojtaba Sharifzadeh	Università di Trento, Italy
Luigi Sibille	Norwegian University of Science and Technology, Norway
Francesco Sica	Sapienza University of Rome, Italy
Adriano Silva	Instituto Politecnico de Bragança, Portugal
Angela Silva	Universidade do Minho, Portugal

Paula Simões	Military Academy, Portuguese Army, Portugal
Francesca Sinatra	University of Trieste, Italy
Mercedes Solla	Universidade de Vigo, Spain
Sergey Solovyev	Sobolev Institute of Mathematics SB RAS, Russia
Maria Somma	University of Naples Federico II, Italy
Federica Sortino	Università degli Studi di Enna Kore, Italy
Eleonora Sottile	University of Cagliari, Italy
Martina Spinelli	Politecnico di Torino, Italy
Elena Stankova	Saint Petersburg State University, Russia
Alexandra Stankulova	Politecnico di Torino, Italy
Giorgia Sugoni	Links Foundation, Italy
Katarína Sýkorová	Matej Bel University, Slovak Republic
Francesco Tajani	Sapienza University of Rome, Italy
Deodato Tapete	Italian Space Agency (ASI), Italy
Rodrigo Tapia-McClung	Centro de Investigación en Ciencias de Información Geoespacial, Mexico
Maria Tartari	CNR, Italy
Daniela Tavano	University of Calabria, Italy
Tarek Teba	University of Portsmouth, UK
Ana Paula Teixeira	Universidade do Minho, Portugal
Wagner Telles	Universidade Federal Fluminense, Brazil
Luis Teodoro	University of Oslo, Norway
Maria Filomena Teodoro	Instituto Superior Técnico, Portugal
Yiota Theodora	National Technical University of Athens, Greece
Elena Todella	University of Parma, Italy
Graca Tomaz	Instituto Politécnico da Guarda, Portugal
Chiara Torre	University of Malta, Malta
Francesca Torrieri	Politecnico di Milano, Italy
Vincenza Torrisi	University of Catania, Italy
Arianna Travaglini	Lucian Blaga University of Sibiu, Romania
Stefanos Tsigdinos	National Technical University of Athens, Greece
Kamila Turečková	Silesian University in Opava, Czech Republic
Elen Twrdy	University of Ljubljana, Slovenia
Toshihiro Uchibayashi	Kyushu University, Japan
Takahiro Ueda	Juntendo University, Japan
Generoso Vaiano	Pegaso University, Italy
Aleksandar Valjarević	University of Belgrade, Serbia
Gianmarco Vanuzzo	University of Perugia, Italy
Dmitrii Vasiunin	Deutsche Telekom Cloud Services E.P.E., Greece
Mohammad Vazir Panah	University of Salerno, Italy
Anna Vaňová	Matej Bel University, Slovak Republic
Enrica Vecchi	University of Cagliari, Italy

Roberto Ventura	Università degli Studi di Brescia, Italy
Ferdinando Verardi	Pegaso Telematic University, Italy
Clara Stella Vicari Aversa	Università degli Studi Mediterranea di Reggio Calabria, Italy
Gianluca Vinti	Università degli Studi di Perugia, Italy
Katarína Vitálišová	Matej Bel University, Slovak Republic
Weronika Wojtak	CCG/ZGDV Institute, Portugal
Fukuko Yuasa	KEK, Japan
Madina Yussubaliyeva	D. Serikbayev East Kazakhstan Technical University, Kazakhstan
Paola Zamperlin	University of Florence, Italy
Francisco Zayas Gato	Universidade da Coruña, Spain
Nodir Zaynalov	Samarkand Branch of Tashkent University of Information Technologies Named After Muhammad Al Khwarizmi, Uzbekistan
Ljiljana Zivkovic	Republic Geodetic Authority, Serbia
Piero Zizzania	University of Naples Federico II, Italy
Renato Zona	Università della Campania L. Vanvitelli, Italy
Floriana Zucaro	University of Naples Federico II, Italy
Francesco Zullo	University of L'Aquila, Italy
Giovanni Zurlo	INFN CNAF, Italy
Özge Öztürk Hacar	Yildiz Technical University, Turkey
Isidora Đurić	University of Novi Sad, Serbia

Plenary Lectures

Digital Twins for Cities – The Rise of the Socio-Technical System

Claudia van der Laag

Department of Built Environment, Oslo Metropolitan University, Norway

Abstract. Cities are entering a new era in which digital twins are no longer simply technical replicas of urban systems, but evolving socio-technical infrastructures that shape how cities are planned, governed, and experienced. Across the world, urban digital twins are being adopted to address some of the most pressing challenges of our time — climate adaptation, urban resilience, mobility transitions, decarbonisation, infrastructure management, and inclusive governance. Yet as these technologies become increasingly sophisticated, a fundamental question emerges: can digital twins truly support more democratic and human-centred cities, or will they reinforce existing structures of power and expertise? In this keynote, Claudia van der Laag explores the transformation of urban digital twins from data-driven modelling tools into collaborative platforms for urban decision-making and societal change. It argues that the future of digital twins lies not only in advances in artificial intelligence, real-time data integration, simulation, and predictive analytics, but equally in the ability to connect technological innovation with governance, participation, and human agency. Drawing on emerging international research and practical case studies, the keynote examines the complex ecosystem surrounding urban digital twins, bringing together planners, policymakers, researchers, technologists, and communities. It highlights how digital twins increasingly function as meeting points between multiple forms of knowledge — technical, political, social, and experiential — and why interdisciplinary collaboration is essential for their success. The talk critically reflects on the tensions between technological ambition and civic inclusion, exploring themes such as transparency, interoperability, participation, and trust. It also discusses how digital twins can become powerful environments for scenario exploration, urban experimentation, and collective learning, enabling cities to better navigate uncertainty and complexity in an age of rapid urban transformation. Ultimately, this keynote presents a vision of digital twins not as neutral technological systems, but as living socio-technical infrastructures capable of reshaping urban governance itself. The challenge ahead is therefore not only how accurately cities can be modelled, but how inclusively they can be imagined, negotiated, and governed through these emerging digital environments.

Prescriptive AI in Practice: Bridging Optimisation, Generative AI and Agentic Decision-Making

Bernardo Almada-Lobo

Universidade do Porto, Faculdade de Engenharia, INESC TEC, Portugal

Abstract. Prescriptive AI, viewed through the lens of mathematical optimisation and operations research, addresses complex business decision-making problems involving limited resources, multiple decision options and competing objectives. This talk considers the full decision pipeline: problem framing, model formulation, computational solution, implementation, monitoring and organisational adoption. The practical success of prescriptive AI depends less on algorithmic sophistication alone than on the quality of the translation between real-world decision contexts and formal optimisation models. In practice, this translation remains a major bottleneck. Objectives are often ambiguous, constraints are incomplete or tacit, data may be fragmented, and the operational consequences of a recommendation may be poorly understood. As a result, optimisation projects can fail even when the underlying mathematical model is technically sound. The talk will address the persistent gap between optimisation theory and deployment practice. While optimisation has generated substantial value in domains such as supply chain planning, logistics, production, revenue management, workforce planning and inventory control, many initiatives still face a significant risk of under-delivery. Common failure modes include incorrect problem framing, weak data readiness, brittle assumptions, infeasible or non-actionable recommendations, limited explainability, insufficient ownership by business users and inadequate post-deployment monitoring. Closing this gap requires moving from one-off model building to an operational lifecycle in which models are stress-tested, embedded in workflows, monitored for drift and revised as the business environment changes. Recent developments in generative and agentic AI create new opportunities to reduce these barriers. Generative AI can support the elicitation of decision-maker preferences, the structuring of unstructured operational knowledge, the drafting of candidate formulations and the explanation of model outputs in business language. Agentic AI can extend this role by orchestrating multi-step workflows that involve model formulation, solver-code generation, execution, debugging, infeasibility diagnosis, sensitivity analysis and human review. However, these capabilities should not be interpreted as replacing optimisation expertise. Rather, they create a new interface between human judgement, analytical modelling and computational decision support. This talk will discuss a cross-fertilisation agenda between optimisation, generative AI and agentic AI, with a focus on how these technologies should be calibrated to the decision at hand.

Net-Zero Communities through Life Cycle Thinking: Canadian Lessons and Global Implications

Kasun Hewage

School of Engineering (Okanagan Campus), Clean Energy Research Centre (CERC), University of British Columbia (UBC), Canada

Abstract. This keynote presentation discusses how life cycle thinking supports the transformation of communities toward sustainable, resilient, and net-zero futures. Drawing on Canadian research, practice, and policy experience, the presentation highlights how integrated life cycle approaches can move beyond conventional carbon accounting to inform practical decision-making across buildings, infrastructure, energy systems, waste management, water systems, and industrial processes [1–8]. The keynote features research conducted at the Life Cycle Management Laboratory (LCML) at the University of British Columbia, Canada, an internationally recognized research hub led by Kasun Hewage. LCML brings together engineers, scientists, policy researchers, postdoctoral fellows, visiting scholars, and graduate researchers to translate data-intensive sustainability research into actionable climate solutions. The laboratory's work is organized around six interconnected research themes: energy systems, urban development, water systems, waste management, construction and asset management, and industrial processes and products. Together, these themes provide an integrated platform for advancing life cycle–based pathways to net-zero communities. Supported by strong partnerships with government, industry, and community stakeholders, LCML has contributed to a substantial body of applied sustainability research, including more than 400 journal articles, 100 conference papers, and numerous technical reports that have informed real-world planning and implementation. This keynote presentation highlights selected success stories that demonstrate how life cycle thinking can be translated into practical strategies for reducing environmental impacts while strengthening community resilience. These insights are especially relevant to the interdisciplinary focus of ICCSA 2026.

References

1. Weerasinghe, O., Hewage, K., Razi, F., Fatmi, M., Sadiq, R.: Applicability of land use models for sustainable urban planning: a review. J. Urban Plann. Dev. ASCE, **151**(3) (2025). https://doi:10.1061/JUPDDM.UPENG-5705

2. Sultana, S. R., et al.: Looking beyond energy efficiency targets: life cycle sustainability of mechanical and water heating equipment in Canadian homes. J. Clean. Prod. Elsevier, **473**, 143469 (2024). https://doi.org/10.1016/j.jclepro.2024.143469
3. Liyanage, D. R., Hewage, K., Hussain, S. H., Razi, F., Sadiq, R.: Climate adaptation of existing buildings: a critical review on planning energy retrofit strategies for future climate. Renew. Sustain. Energy Rev. Elsevier, **199**, 114476 (2024). https://doi.org/10.1016/j.rser.2024.114476
4. Liyanage, D. R., Hewage, K., Karunathilake, H., Sadiq, R.: Feasibility of carbon-capturing in building heating systems: a life cycle thinking-based approach. International J. Greenhouse Gas Control. Elsevier, **132** (2023). https://doi.org/10.1016/j.ijggc.2023.104056
5. Moudgil, V., Sadiq, R., Bakhtavar, E., Paudel, A., Hewage, K.: A cloud-oriented data-analysis framework to analyze peak demand dynamics in institutional building clusters. Sustain. Cities Soci. Elsevier, **111**, 105553 (2024). https://doi.org/10.1016/j.scs.2024.105553
6. Liyanage, D. R., Hewage, K., Ghobadi, M., Sadiq, R.: Thermal resiliency of single-family housing stock under extreme hot and cold conditions. Energy Build. Elsevier, **323**, 114809 (2024). https://doi.org/10.1016/j.enbuild.2024.114809
7. Wijayasekera, S. C., et al.: Sustainability of waste-to-hydrogen conversion pathways: a life cycle thinking-based assessment. Energy Convers. Manage. Elsevier, **270** (2022). https://doi.org/10.1016/j.enconman.2022.116218
8. Jayasena, A., Hewage, K., Siddiqui, O., Sadiq, R.: Socio-economic and environmental cost-benefit analysis of passive houses: a life cycle perspective. J. Clean. Prod. Elsevier, **373** (2022). https://doi.org/10.1016/j.jclepro.2022.133718

Machine Learning Based Methodologies for Grape and Wine Production

Pedro Melo-Pinto

University of Trás-os-Montes and Alto Douro, Engineering Department, Centre for the Research and Technology of Agro-Environmental and Biological Sciences, Portugal

Abstract. Wine production is a major area in the Portuguese economy. The sector accounts for approximately 2.6% of Portugal's GDP, and is crucial to the development of inland regions. Wine quality depends on grape quality, which is a result of different factors such as harvesting at the optimal point of maturity, or selecting the grapes according to quality features. The growing importance of precision agriculture demands novel methods to make the most out of the data generated. Machine learning approaches are particularly suitable for modelling complex processes, where too many arguments are involved. Many of the agricultural processes such as yield estimation or ripening fall into this category. However, several challenges may arise: even if models are available for particular cases, general models are difficult to come by, largely due to the high variability involved; with such models, there is a need for huge amounts of data, in a way which will not always be easy to obtain. This is particularly true in the case of grape production and the wine industry, where the number of grape varieties, climate, soil, etc., introduce a degree of variability that is uncommon in other sectors. One way of obtaining more information in such cases has been the use of images, whether RGB, multi-spectral, or hyperspectral. The use of images introduces additional challenges, such as data size and lighting conditions, which underscores the need for data-driven modelling methodologies. The use of machine learning (ML) techniques combined with images has been establishing itself as a viable alternative for predicting yields or quality parameters in grapes, and assisting with critical harvesting decisions. Several ML approaches have been proposed to handle such data characteristics, but selecting a suitable methodology that best addresses the problem under study and making sure it generalizes well is, most of the time, a cumbersome task. We will address three wine and grape production problems, yield estimation, grape quality assessment, and illness detection, presenting different difficulties. Yield estimation is a major aspect in the wine production industry. The number of berries is one of the most relevant yield aspects, and can be estimated using computer vision, namely image segmentation algorithms. Lighting is one of the key factors in computer vision approaches, and occlusions—which are common in this problem—can be minimized, but require estimation models to achieve reliable results. Early illness detection is, of course, of the utmost importance, with an impact on production and quality. Detecting early symptoms and distinguishing them from other conditions is a difficult task, which makes the use of machine learning algorithms advisable, using automatic visual inspection of the leaves, addressed as a pixel-based classification problem. However, the available datasets are limited and

complex to compile, as different diseases typically require very specific time-frames and locations. The use of generative models to overcome this obstacle, in conjunction with robust classification models, leads to models that are intended to be as general as possible. The evaluation of the grape's maturation profile based on the evolution of their oenological parameters over time is one of the most important aspects for ensuring wine quality, playing a decisive role in the definition of the optimal time for harvesting. The use of hyperspectral imaging combined with learning algorithms represents a faster and non-destructive grape ripeness assessment method. We focus in two fundamental aspects to address the resulting data variability: a possible wavelength bands selection (with the purpose of reducing the dimensionality of data without losing predictive power) and the generalization ability of the ML model under such demanding conditions.

Contents

Modelling Liveable Cities: Techniques, Methods, Challenges, and Perspectives Behind the X-Minute City (MLC 2026)

From Fragments to Systems: Mapping Urban Spatial Potential for X-Minute City Planning 3
Chiara Anselmi, Camilla Pezzica, and Federica Deri

Distribution of Spatial Inequalities Across Social Groups: a Case Study of Messina, Italy 22
Beniamino Murgante and Alfonso Annunziata

The Limits of the 15-Minute City: a Systematic Qualitative Literature Review 39
Beniamino Murgante and Alfonso Annunziata

Main Streets or Latent Corridors? Exploring the Significance of Weak Ties in Suburban Contexts 60
Federico Mara, Federica Deri, and Valerio Cutini

Mathematical Methods for Image Processing and Understanding (MMIPU 2026)

On Xylographic Image Reconstruction: the Segmentation Phase 77
Ivan Gerace

The Minimum Sum Coloring Problem on Circulant Graphs $C_n(1, 2)$ 95
Eugenio Maria Cocco and Sara Nicoloso

A Smoothing Aggregation Cascadic Multilevel Method for Image Deblurring 108
Pietro Maurino and Marco Donatelli

Neural Network and Sampling Kantorovich Operators vs Traditional Denoising Techniques 124
Ana Maria Acu, Cristian Daniel Alecsa, Arianna Travaglini, and Gianluca Vinti

Multivariate Generalized Gamma-Type Operators: Approximation Results and Applications to Image Processing 142
Vita Leonessa, Arianna Travaglini, and Gianluca Vinti

Parameters Estimation via Unrolling for the Preconditioned Nested Primal Dual Algorithm 156
Marco Donatelli and Giuseppe Scarlato

Numerical Aspects in Machine Learning For Dynamical Systems (NAMLYS 2026)

Intrinsic and Post-hoc Interpretability of Gaussian-Based Kolmogorov–Arnold Networks 173
Pasquale De Luca, Emanuel Di Nardo, Emilio Garzia, Livia Marcellino, and Angelo Ciaramella

A GPU Accelerated DCT Implementation for Image Compression 184
Angelamaria Cardone and Gerardo Di Pascale

Numerical Modeling for Information Diffusion on Social Media: a Review of Recent Works 199
Angelamaria Cardone, Dajana Conte, Patricia Díaz de Alba, Samira Iscaro, Giovanni Pagano, and Beatrice Paternoster

Downsampling Strategies for LES Data from High-Fidelity CFD Solvers 221
Raffaella Sapio, Carmine De Lucia, Andrea Mastellone, Alfonso Matrone, Gerardo Toraldo, and Ferdinando Zullo

Numerical Methods and Machine Learning in Science & Engineering (NMLSE 2026)

A Comparative Study of Optimization Models for Superellipsoid Fitting in 3D Object 241
Maria Silva, Fernanda Costa, and Eliana Costa e Silva

On the Dependency of Neural Differential Equations on the Training Numerical Method: Effects on Optimisation and Predictive Performance 255
C. Coelho and O. Niggemann

Recurrent Neural Networks for Predicting Inflows in Hydroelectric Reservoirs 272
E. Villar, Cecília Coelho, M. Fernanda P. Costa, and L. L. Ferrás

Parameter Estimation for a Springpot–Fractional Dashpot Viscoelastic Model: Optimisation and Numerical Laplace Inversion 285
L. L. Ferrás, C. Coelho, M. Fernanda P. Costa, M. L. Morgado, and M. Rebelo

Chaotic System Fingerprinting for Cryptographic Keystream Attribution: A Deep Learning Approach 298
Alexandru Dinu, Sara-Gabriela Modan, and Mădălin Frunzete

Physics-Informed Neural Networks for High-Dimensional Black–Scholes Option Pricing: Accuracy, Scalability, and Limitations 313
André Paz Gomes, Isaac P. Santos, and Lucia Catabriga

6th Workshop on: Privacy in the Cloud/Edge/IoT World (PCEIoT 2026)

A Lightweight CRYSTALS-Dilithium Post-quantum Cryptography Signature for FPGA-Based Edge Computing 333
Huy Minh Dung Do, Duc Hoang Pham, Thien Hai Ha, and Cuong Pham-Quoc

A Binning-Based Approach for Intrusion Detection and Attack Classification in Internet of Medical Things Networks 351
Lerina Aversano, Stefano Galantucci, Michele Mastroianni, and Andrea Porcelli

Preserving Our Past: Spatial and Remote Sensing Technologies for Cultural Heritage in a Changing Climate (POP 2026)

From Open Land to Built Surfaces: Spatial Analysis and Remote Sensing for Photovoltaic Energy Computing 371
Maria Danese and Canio Alfieri Sabia

Spatial Patterns of Urban Air Pollution Exposure in Cultural Heritage Areas Derived from Downscaled Satellite Observations 385
Maria Danese, Felice Perciante, and Rosa Lasaponara

Scientific Computing Infrastructure (SCI 2026)

Optimizing Harmful Information Containment in Social Networks via Shortest-Path Interdiction 403
Muon Ha, Yulia Shichkina, Tran Duc Manh, Xuan -Hien Nguyen, and Thanh-Son Phan

New Digital Birth Certificate Based on Biometric and Documentary Data of the Child 415
A. D. Shchegolev, G. A. Kukharev, and N. L. Shchegoleva

Identifying the Characteristics of Brain Regions in Individuals with Different Neurocognitive Characteristics Based on Functional MRI Data 433
Nadezhda Shchegoleva and Kakabayeva Nurjemal

TFPACO: Traffic Flow Planning System for Software-Defined Networks Based on Ant Colony Optimization Algorithm 447
Aleksandr Safonenko and Vladimir Korkhov

Multithreaded Pipeline Parallelism for Distributed LLM Inference Without Source Code Modification 464
Victor Koloskov and Anton Pershin

Multilayer Computing Systems 481
Evgeniy Ibatullin and Alexander Bogdanov

Avg-Frontier: A Low-Overhead Multi-objective Batch Scheduler for Heterogeneous Edge Environments 493
Aleksei Chernigov and Vladimir Korkhov

Creating an Operating System for Video Camera Chips in an Intelligent Video Surveillance System 511
Alexey Subbotin, Nataly Zhukova, and Elena Stankova

Sustainable evolution of long-Distance frEight and paSsenger Transport (SOLIDEST 2026)

Comparison of Container Routing in Liner Shipping Networks 527
Massimo Di Gangi, Orlando M. Belcore, and Antonio Polimeni

Analysis and Interpretation of Attributes Affecting Ship Time in Container Ports 543
Giuseppe Musolino, Paola Panuccio, Corrado Rindone, Francesco Russo, and Antonino Vitetta

Modal Choice for High-Speed Railway Systems: The Italian Case Study 556
Giuseppe Musolino, Marialuisa Moschella, Francesco Russo, and Domenico Sgro

Environmental Sustainability of Railway Corridors: Case Studies of Italian High Speed Rail 570
Giuseppe Musolino, Marialuisa Moschella, Francesco Russo, and Domenico Sgro

An Assessment of Road Network Topology and Its Influence on Port Accessibility in Romania 585
Aura Rusca, Florin Rusca, Alina Roman, Oana Dinu, Stefan Burciu, Iulia Dorobantu, and Denis Codroiu

A LEAN Approach for High-Speed Railways: General Elements for Extending the Italian Network in the South 598
Corrado Rindone and Antonio Russo

Sustainable Tourism Economy, Silk Road, Reuse and Enhancement of Cultural Heritage (STESilkRoad 2026)

Cultural Heritage, Identity, and Power Dynamics: Guiding Practices in Oradea, Romania 615
Corina-Florina Tătar, Celestina Fazia, Clara Stella Vicari Aversa, and Klara Szylagyine Czimre

Efficient Use of Natural Tourism Resources Within Sustainable Tourism Along the Silk Road: Forecasting Tourist Flows in Uzbekistan Based on the ARIMA Model 630
Bahodirhon Safarov, Mansurova Nargiza, and Mustafokhon Toshkhujaev

Author Index 643

Modelling Liveable Cities: Techniques, Methods, Challenges, and Perspectives Behind the X-Minute City (MLC 2026)

From Fragments to Systems: Mapping Urban Spatial Potential for X-Minute City Planning

Chiara Anselmi(✉), Camilla Pezzica, and Federica Deri

Department of Energy, Systems, Territory and Construction Engineering, University of Pisa, Largo Lucio Lazzarino, 56122 Pisa, Italy
chiara.anselmi@ing.unipi.it

Abstract. In contemporary urban planning, fragmented suburban contexts and limited public resources require operational tools able to identify where public spaces and facilities can be introduced with limited conflict, cost, and land consumption. This paper proposes a parcel-based computational decision-support framework for mapping urban spatial potential in X-minute city planning. Taking cadastral parcels as the elementary unit of transformation, the model evaluates land suitability for two distinct targets – open public spaces and built public facilities – through a multi-criteria structure addressing morphology, regulatory framework, ownership and accessibility issues, lot coverage and uses, environmental and landscape constraints. The framework is applied to the suburban area of Massa (Italy), a heterogeneous territory characterized by strong spatial fragmentation, low service provision, and high dependence on private mobility. These conditions make it a particularly relevant testing ground for proximity-based planning approaches, where the challenge lies in activating and interconnecting dispersed spatial resources. Results highlight differentiated spatial potentials for public spaces and facilities, confirming that proximity planning cannot rely on a single transformation logic. At the same time, private interests and ownership emerge as critical structural constraints, emphasizing the need to explicitly integrate property conditions into computational planning tools. The study also outlines how parcel-level suitability can support a network-based interpretation of urban systems, enabling the identification of strategic nodes and connections. This work provides a foundation for future research aimed at adapting the model to different policy scenarios and integrating advanced network analytics, supporting more flexible and context-sensitive implementations of proximity-based planning.

Keywords: Urban Spaces · Parcel-based Approach · 15-min city

1 Introduction

Ensuring equitable access to public spaces, services, and everyday opportunities has become a central concern in contemporary urban planning, particularly in contexts shaped by dispersed urban growth and increasing pressure on public resources. Within this broader framework, a persistent challenge is the fragmentation of public spaces, which often appear as isolated "islands" within the urban fabric [1]. These disconnected

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 3–21, 2027.
https://doi.org/10.1007/978-3-032-30530-5_1

areas suffer from severe pedestrian and vehicular accessibility issues, often compounded by parking shortages [2]. At the same time, the debate on public space provision has highlighted the inadequacy of traditional minimum planning standards, which in practice often reduce public amenities to residual green areas or parking lots within private subdivision plans, without any broader city-wide vision [3, 4].

In response to these spatial inequalities, proximity-based planning and the "15-min city" have gained increasing attention [5]. This paradigm is particularly relevant in the suburban contexts of Italian cities, which are often characterized by residential monofunctionality and a strong dependence on private mobility for access to essential services [6, 7]. Within this perspective, the emc2 model provides a useful reference, proposing the transformation of main streets in low-density suburban areas into linear pedestrian corridors enriched by a mix of public and private services, with the aim of enhancing accessibility and revitalizing peripheral neighborhoods [8].

Despite the theoretical appeal of proximity models, a significant research gap remains in the operational tools available for their implementation. Municipalities are increasingly constrained by limited public resources and frequent conflicts with private stakeholders [9, 10]. As a result, implementing the proximity city is not only a morphological issue, but also a matter of property and ownership. Effective local planning cannot ignore existing land uses and ownership structures, since neighborhood-scale regeneration must be grounded in the socio-economic reality of land. This perspective is consistent with the interpretation of cities as complex adaptive systems rather than as products of top-down design [11, 12]. Urban form emerges through the aggregation of localized decisions, many of which take place at the scale of the individual cadastral parcel, which becomes a crucial operational unit, enabling planners to mediate between private rights and collective interests under real spatial and economic constraints [13, 14].

Building on these premises, this study operationalizes the concept of "urban vocation" through a parcel-based computational decision-support model. Rather than prescribing fixed outcomes, the model evaluates current land profiles in order to identify parcels whose existing morphological, regulatory, environmental, accessibility, and ownership conditions make them more suitable for future public transformation. By incorporating legal and economic constraints, the framework supports the identification of opportunities for public intervention while reducing costs and potential land-use conflicts, thus linking bottom-up urban dynamics with strategic planning choices [15].

At the same time, operationalizing the emc2 model requires transformed parcels to function not as isolated opportunities, but as parts of a connected spatial system supporting accessibility, continuity, and daily mobility. For this reason, the study complements parcel-level suitability with a network-based perspective that explores how suitable parcels can relate to one another within a wider proximity system. Relying on network-based measures such as shortest paths or configurational metrics [16, 17], the proposed approach supports the identification of candidate nodes, the evaluation of their connectivity, and the exploration of scenarios capable of maximizing spatial coverage while limiting intervention costs. This ultimately translates the "latent potential" of parcels into operable spatial strategies.

The methodology is tested on the case study of Massa, Italy, a sprawling suburban area between the mountains and the coastline, characterized by fragmented settlement patterns and a marked deficit of public spaces and services.

The remainder of this paper is organized as follows. Section 2 discusses the main issues related to cadastral datasets. Section 3 presents the methodological framework. Section 4 describes the case study area. Section 5 reports and discusses the results of the suitability analysis, and introduces a preliminary facility-facility graph. Finally, Sect. 6 concludes by summarizing the main findings and outlining future developments.

2 Dataset-Related Issues

The aim of this contribution is to engage with local-level implementation planning by bringing public regulations into alignment with private action and its management. This requires a fine-grained analysis of urban form, in which the cadastral parcel is treated as the actual decision-making unit of urban transformation.

Recent studies have demonstrated that plot configurations play a decisive role in shaping patterns of physical change, as they constitute one of the main determinants of urban form and actively guide its evolution [18, 19]. Within parcels, form, ownership, land use, and planning rules converge as structural conditions of urban action [20]. Parcels therefore become a policy instrument aimed at balancing individual property rights with collective environmental and spatial well-being.

Although cadastral data represent the official legal reference for taxation and regulatory constraints, they are often outdated and spatially inconsistent with other topographic and open datasets. These discrepancies may result in positional inaccuracies that affect both administrative processes and planning decisions [21–23]. To address this issue, the proposed model does not rely on strict spatial overlaps between datasets. Instead, it introduces tolerance thresholds and rule-based tolerances for each indicator, allowing for a more flexible and robust integration of heterogeneous data sources. This results in a quasi-fuzzy evaluation framework, in which suitability is derived from degrees of proximity and compatibility rather than exact geometric coincidence. Moreover, cadastral parcels do not always reflect the current physical configuration of the urban environment, particularly where land subdivision has already occurred to accommodate infrastructures such as roads, without corresponding updates in cadastral classification or geometry. To resolve this mismatch, the cadastral layer has been reinterpreted through its integration with additional datasets and orthophotography.

Furthermore, the analysis focuses exclusively on parcels within the urban fabric that are realistically susceptible to transformation. Accordingly, the original dataset has been refined by excluding roads, water bodies, and already structured natural elements, which do not constitute ordinary transformation units. Existing public spaces (such as streets, squares, parks, and service facilities, as defined by the Municipal Urban Plan of Massa) are also excluded from the pool of candidate parcels. However, they are not removed from the model; rather, they are retained as structural components of the system and as essential benchmarks for the subsequent network analysis.

3 Methods

The proposed decision-support model focuses on identifying suitable locations for public spaces and public facilities in suburban areas, where such provisions are often limited (see Sect. 4). The model distinguishes between transformations into open public spaces and those into built public services, as these two categories respond to fundamentally different spatial and functional logics. While open public spaces are flexible and multifunctional, supporting both social and environmental roles, built public facilities require full compliance with building and infrastructural conditions, including direct accessibility, service connections, and regulatory constraints, resulting in a more limited and structure-dependent locational potential. Within the category of open public spaces, the model considers the allocation of collective outdoor activities, such as public squares and urban parks. Areas explicitly designated for ecological corridors are not treated as a separate output category. Instead, urban greenery is reinterpreted not merely as a residual or standard-based provision, but as a structurally integrated system. In this sense, urban green spaces are conceived as interconnected elements capable of enhancing accessibility and the quality of public space, while simultaneously supporting ecological functions.

The resulting suitability model captures the inherent predisposition of land to accommodate public transformations, avoiding a deterministic interpretation that may fail to capture unfavourable local conditions, potentially affecting the feasibility of the resulting interventions. It adopts a compromise-based approach, in which no configuration is strictly excluded from the set of possible transformations. Lower suitability values indicate increasing levels of implementation complexity, higher costs, or the need for regulatory adjustments, rather than infeasibility. This reflects the incremental and adaptive nature of urban transformation processes, where multiple suboptimal conditions can still be negotiated within a broader decision-making framework.

3.1 Suitability Analysis Criteria

The multicriteria decision-making model is based on a set of eight criteria capturing the main spatial conditions influencing parcel transformation. These criteria are designed to represent a different dimension of urban feasibility, including:

- morphological characteristics,
- accessibility and network integration,
- land ownership conditions,
- urban context,
- regulatory constraints,
- environmental factors.

Each criterion is articulated through a set of descriptors, which are subsequently translated into quantitative parameters defining thresholds and spatial relationships (see Table 1). Together, they provide a multi-dimensional representation of the conditions that enable or constrain the transformation potential of each parcel. In the following paragraphs each criterion will be analysed.

Morphological Development Potential. This criterion assesses whether parcel form is compatible with feasible urban transformation. Since spatial constraints directly affect

costs and implementation complexity, small, narrow, or irregular plots are generally less adaptable than larger and more regular ones [18, 19]. Feasibility is operationalised through a minimum buildable footprint of 300 m2, calculated after applying a 10-m setback from buildings in accordance with Italian regulations concerning windowed façades (Art. 9 DM 1444/68). Distances from roads and parcel boundaries are not considered fixed constraints, as they can be adjusted within implementation-level planning processes. Additionally, a compactness index (I_c calculated from area and perimeter, following the normalised formulation in Eq. 1) is used to capture the geometric regularity of parcels, with a threshold of 0.25 excluding highly elongated or irregular plots that are less compatible with efficient and standard development conditions [24, 25].

$$I_c = \frac{4\pi A}{P^2} \tag{1}$$

Regulatory Development Potential. Beyond morphological constraints, development potential is also shaped by regulatory restrictions unrelated to parcel form. The model therefore includes buffer zones associated with major transport infrastructures, such as highways and railways (30 m), as well as other locally regulated areas and site-specific protected zones, including cemeteries, protected green areas, and DOC/DOP agricultural land. While these conditions may considerably limit private development, they can, where compatible, open opportunities for public intervention.

Land Ownership. Adopting a realistic approach, ownership is treated as a structural condition of urban change rather than as a simple contextual attribute, since it determines decision-making power and the actual possibility of transformation [20, 26]. Given the financial limits of public land acquisition and the widespread fragmentation of private property, publicly owned parcels represent a strategic starting point for expanding a continuous and interconnected public urban network [27].

Road Network Accessibility. This criterion evaluates the relationship between parcels and the existing mobility network, assuming accessibility to be a necessary condition for effective public intervention. Proximity to roads improves usability and reduces the need for additional infrastructure, although excessive proximity may create safety or environmental concerns for recreational and ecological uses [28–31]. In this study, accessibility is prioritised: parcels with direct or near-direct access receive higher suitability scores, with a 5-m tolerance for geometric inconsistencies and distances beyond 100-m considered poorly accessible.

Lot Coverage. Lot coverage is used as a proxy for settlement saturation and, more broadly, for the degree of soil sealing and land artificialisation (LR Toscana 65/2014). In line with the objective of limiting further land consumption, already developed areas are generally considered more suitable for reuse, although no distinction is made here between different types of built occupation. A reference threshold of approximately 30% is adopted as a benchmark for high land occupation (S_c), with lower values generally associated with higher transformation potential.

Land Use. Land use influences both the feasibility and the appropriateness of intervention, as it reflects existing development conditions, ecological value, and potential acquisition costs. Already artificialised areas are generally more suitable for new public

facilities, whereas agricultural, semi-natural, and natural areas may require preservation, careful management, or lower-impact forms of collective use. Parcels are therefore classified into four categories—artificial surfaces, agricultural land, semi-natural green spaces, and natural areas—representing a gradient of transformation potential.

Environmental Constraints. This criterion incorporates the most restrictive and generally non-derogable conditions affecting land transformation, including flood hazard (PGRA), geological hazard (PAI), and protected areas such as the Natura 2000 network. Because these constraints impose significant limits on land-use change, especially in high-risk areas, parcels subject to them show substantially lower transformation potential, particularly for built interventions.

Landscape and Heritage Constraints. Site-specific conditions related to landscape quality, cultural value, and spatial identity are addressed through the framework of D.Lgs. 42/2004 (*Codice dei beni culturali e del paesaggio*). These include buffers around lakes, rivers, and coastal areas, as well as forests, woodlands, parks, and listed heritage buildings. Unlike environmental constraints, they do not necessarily prevent transformation, but rather condition its form, scale, and compatibility with existing landscape and cultural values. Parcels not affected by constraints are assigned the highest suitability.

3.2 Suitability Analysis Workflow

The model adopts a multi-criteria aggregation procedure to derive parcel-level suitability values from the previously defined criteria. As mentioned above, rather than defining strict suitability thresholds, the approach is designed to capture varying degrees of feasibility, reflecting the incremental and context-dependent nature of urban transformation processes. Each descriptor within a given criterion is assigned a score on a scale ranging from 0.25 to 1, according to its contribution to the suitability for a specific target use. Lower values indicate increasing levels of implementation complexity, cost, or required intervention. No criterion is treated as a strict constraint and no score equal to zero is assigned. Table 1 shows the scoring assessment for all the criteria.

Although each criterion is evaluated independently, their effects on urban transformation are not independent. The feasibility of intervention depends on the combined presence of multiple conditions rather than on their additive contribution. The aggregation is performed through a geometric mean (2) that avoids full compensability between criteria and preserves sensitivity to low-performing conditions without excluding transformation entirely. This results in a compromise-based suitability structure in which all criteria contribute jointly, without generating exclusionary effects. The formulation is applied independently to each target use, namely public space and public facilities, producing distinct suitability outputs. No additional weighting scheme is introduced, as all criteria are conceptualised as "should-have" conditions and are therefore treated as equally contributing factors.

This ensures a non-hierarchical aggregation structure and avoids the introduction of arbitrary prioritisation among criteria. To complement the aggregated suitability score, the model retains information on the most limiting condition affecting each parcel. While the geometric mean provides a synthetic measure of overall suitability, it does not

Table 1. Suitability scoring assignment for public spaces and facilities

Criteria	Descriptors	Parameters	Public Space Score	Public Facilities Score
Morphological Development Potential	High development suitability	$A \geq 300$ mq, $Ic \geq 0.25$	0.25	1.00
	Moderate development suitability	$A \geq 300$ mq, $Ic < 0.25$	0.50	0.75
	Low development suitability	$A < 300$ mq, $Ic \geq 0.25$	0.75	0.50
	Unsuitable configuration	$A < 300$ mq, $Ic < 0.25$	1.00	0.25
Regulatory Development Potential	No constraints	overlap = 0%	0.25	1.00
	Partially constrained	overlap < 40%	0.50	0.75
	Significantly constrained	40 < overlap ≤ 90%	0.75	0.50
	Fully constrained	overlap > 90%	1.00	0.25
Land Ownership	Private	overlap > 70%	0.25	0.25
	Religious institutions	overlap > 70%	0.50	0.50
	Local government	overlap > 70%	0.75	0.75
	Publicly owned	overlap > 70%	1.00	1.00
Road Network Accessibility	Direct access	$0 < d \leq 5$ m	1.00	1.00
	Limited accessibility	$5 < d \leq 100$ m	0.625	0.625
	Poor accessibility	$d > 100$ m	0.25	0.25
Lot Coverage	High coverage	$Sc \geq 0.3$	0.25	0.25
	Low coverage	$0 < Sc < 0.3$	0.625	0.625
	Vacant lot	$Sc = 0$	1.00	1.00
Land Use	Artificial surfaces	Main use	0.25	1.00
	Agricultural land	Main use	0.50	0.75
	Semi-natural green spaces	Main use	0.75	0.50
	Natural areas	Main use	1.00	0.25
Environmental Constraints	No constraints	(P1, P2, P3, P4) = 0	0.25	1.00

(*continued*)

Table 1. (*continued*)

Criteria	Descriptors	Parameters	Public Space Score	Public Facilities Score
	Low to medium-low hazard	(P4, P3) < 20% or (P2) < 40% or (P1)	0.50	0.75
	Medium-high to high hazard	20 < (P4, P3) < 40% or (P2) > 40%	0.75	0.50
	Severe hazard	(P4, P3) ≥ 40%	1.00	0.25
Landscape and Heritage Constraints	No constraints	overlap = 0%	0.25	1.00
	Partially constrained	overlap < 40%	0.50	0.75
	Significantly constrained	40 < overlap ≤ 90%	0.75	0.50
	Fully constrained	overlap > 90%	1.00	0.25

explicitly identify which criterion is responsible for score reductions. For this reason, a parcel-level criticality indicator is introduced based on the minimum value among all evaluated criteria (3). This value captures the intensity of the most critical condition, directly indicating the degree to which a specific factor constrains the transformation potential of the parcel.

The formulas are defined as follows:

$$S_p = \left(\prod_{i=1}^{n} c_{i,p}\right)^{\frac{1}{n}} \tag{2}$$

$$C_p = \min_i \left(c_{i,p}\right) \tag{3}$$

where:

- S_p represents the suitability of parcel p,
- C_p represents the lowest-performing criterion for parcel p,
- $c_{i,p}$ is the score of criterion i for parcel p,
- n is the total number of criteria (in this case, 8).

In addition, the model records the identity of the limiting criterion through the index:

$$I_p = \arg\min_i (c_{i,p}) \tag{4}$$

which identifies the criterion responsible for the minimum score.

Finally, the two suitability maps are generated by combining all the defined criteria and their respective scores within the proposed framework. Suitability is represented through a discretised colour scale, where darker shades indicate a higher predisposition

for the allocation of public spaces or services, corresponding to lower implementation costs. The entire workflow has been automated through the QGIS Graphical Modeler, ensuring the reproducibility and scalability of the model.

3.3 Integrating the Suitability Analysis into Proximity Assessments

The parcel-level suitability model identifies where public spaces or facilities can be delivered at lower implementation cost and conflict. Yet proximity (in the X-minute city sense) depends not only on the qualities of individual parcels, but also on their ability to operate as part of a connected spatial system that supports accessibility, continuity, and daily mobility.

This paper, therefore, begins to explore the integration of parcel-level suitability outputs into a network-based proximity model to evaluate spatial relations among facilities (e.g., park-to-school walking continuity) and their coverage. To this end, it proposes the construction of a facility–facility primal graph model [32] that links each candidate parcel to anchor nodes and to other candidates via available street-network edges. In this model, both the node layer and the edge layer represent suitability-driven potential. Candidate nodes are parcels whose suitability for public spaces or public facilities exceeds a given threshold (i.e., 20% highest values). Nodes that are potentially subject to transformation, inherit the suitability score and criticality index from the parcel model, whereas existing public spaces (e.g., schools, parks etc.) represent fixed anchor-nodes. The edge layer represents accessibility-driven feasibility, restricted to connections within a predefined walking-time threshold (namely 15 min). New edges are dedicated pedestrian and cycle paths. Edge costs are computed as shortest-path walking distances or times, optionally weighted by attributes such as sidewalk quality, safety, slope, or permeability. After selecting candidate nodes, and constructing the cost-distance network (with edge weighting), scenario variants – based on different selection logics (e.g., max coverage, minimum cost, ecological) – can be evaluated using decision-oriented planning indicators (e.g., network accessibility, connectivity, redundancy) appropriate to each case.

Importantly, rather than fully applying this model to the case-study analysis, we illustrate the proposal (to be tested in future studies) through an applied example to clarify the model's foundations and discuss its utility for activating parcel-level feasibility (micro-scale) in support of a spatially grounded evaluation of intervention possibilities at the meso-scale, considering both new nodes and links.

4 Case Study

The proposed method has been tested within a selected portion of the Municipality of Massa, extending from the Frigido River to the southern boundary with the Municipality of Montignoso, and from the coastline inland to the hilly areas. Massa is part of the urbanised continuum of the Versilia conurbation. As noted by Mara et al. [33], the progressive merging of the urban fabric across adjacent municipalities has diminished the role of historic centres in favour of broader, system-wide dynamics. However, this process has also reinforced the peripheral condition of suburban areas, concentrating

services within emerging local centres while leaving the interstitial spaces increasingly dependent on private mobility.

The case study is particularly relevant due to the coexistence of multiple urban environments within a relatively limited area. The Frigido River divides the municipality into two distinct parts, while the transversal presence of major infrastructures—namely the highway and the railway—further fragments the territory into sectors with differentiated spatial and functional characteristics. The northern area is predominantly characterised by productive and tourism-related activities, whereas the southern part includes a compact urban centre alongside a suburban residential area marked by dispersed settlement patterns. Between the railway and the motorway, a partially urbanised rural landscape persists, where agricultural uses remain active. In contrast, the area between the motorway and the coastline is characterised by an anthropised pine forest interspersed with low-density residential development, mainly composed of detached houses and villas (Fig. 1). These suburban areas share a common condition of limited provision of public facilities and collective spaces. At the same time, the relatively low level of land consumption makes the area particularly suitable for public-led transformation processes aimed at introducing a coherent network of public spaces and services. As illustrated in the left-hand side of Fig. 1, public spaces and services are predominantly concentrated around the compact urban centre, with a limited number of green or sports-related facilities extending towards the coastal area. A clear deficit of public spaces and services can be observed in the south-eastern portion of the study area, particularly within the sprawling residential areas at the foothills (the Turano neighbourhood) and along the coast (the Ronchi–Poveromo district), as well as within the urbanised rural landscape located between the railway and the highway.

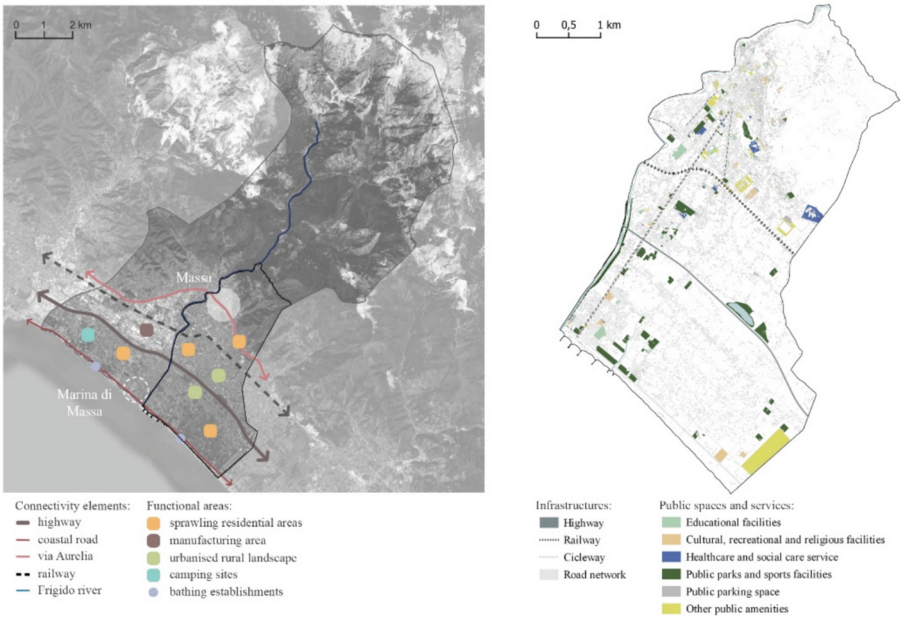


Fig. 1. Functional analysis of the case study area within the Municipality of Massa.

5 Results and Discussion

The results of the suitability analysis reveal a differentiated spatial pattern in the distribution of potential locations for public spaces and facilities. Overall, the model highlights a distinction between the two types of interventions, confirming the underlying assumption that open public spaces and built public services respond to different spatial and functional conditions.

As illustrated in Fig. 2, a strong potential for open public spaces emerges in the foothill belt and, most notably, in the coastal area of Ronchi–Poveromo. Although parcel size and availability might appear favourable for development, these areas are affected by regulatory constraints and land uses that are incompatible with new construction. In particular, the presence of pine forest and woodland in Ronchi–Poveromo, as well as the DOC Candia wine-producing area in the foothill zone, significantly limits buildability while enhancing the suitability for open public uses. Within the compact urban fabric of Marina di Massa, high density and strong cadastral fragmentation also constrain the possibility of new public building development. A particularly relevant case is represented by the axis connecting the historic centre to the coastal settlement of Marina di Massa, where a marked difference between the two suitability maps can be observed. This axis is recognised as a protected visual corridor (“cannocchiale panoramico”), whose visibility is preserved under DCR 82/2022, thus limiting building potential. Similarly, the area surrounding the artificial lake located above the motorway is subject to environmental and landscape constraints that significantly restrict development, while making it particularly suitable for collective open-air use.

In contrast, the suitability for urban public facilities is more evenly distributed across the central lowland belt. In the urbanised rural areas, despite the strong presence of private ownership, the relative absence of constraints and the low level of land occupation favour the introduction of public facilities, including through new construction.

A closer examination of the criteria affecting suitability reveals that, for most parcels, the most recurrent limiting factor identified by the criticality index (I_p) is land ownership. This reflects the predominantly private ownership structure of the area, which represents a major constraint for the implementation of public policies.

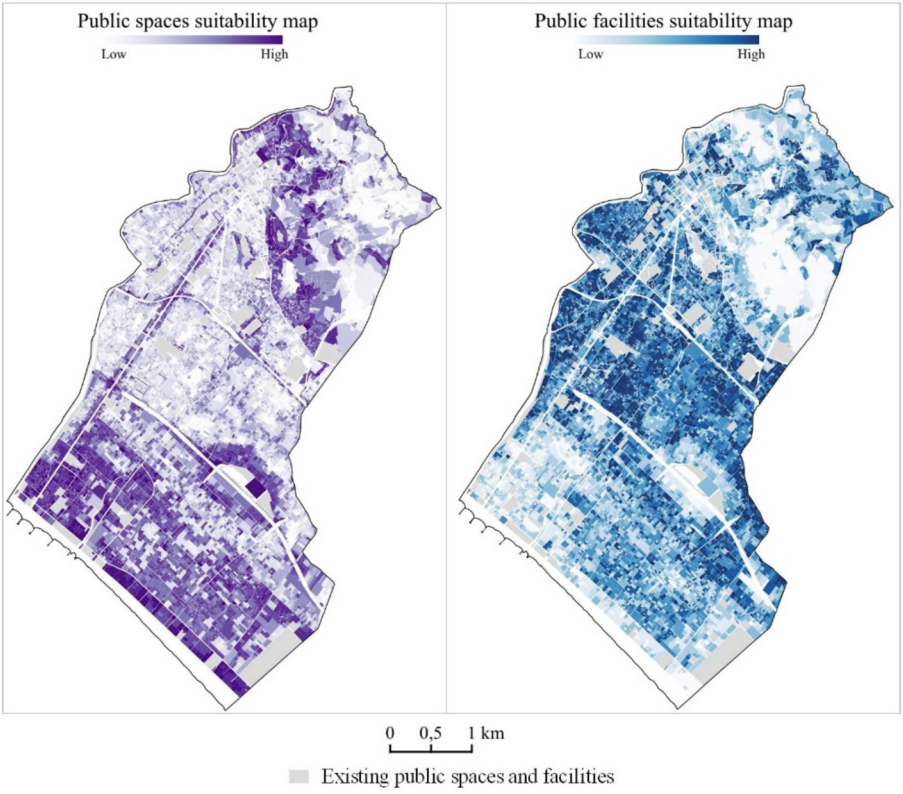


Fig. 2. Suitability maps for public spaces and public facilities.

Due to the fragmented distribution of suitable parcels, considering only those allocated to open public spaces makes it difficult to define a fully continuous and integrated network at the scale of the entire study area. However, local opportunities for spatial continuity can still be identified. In the case of public facilities, spatial continuity is easier to achieve, given their concentration in the central areas, although their presence decreases towards the coastal zone. Ultimately, a coherent and effective network can only be achieved through the integration and overlap of the results of the two suitability maps. Given that parcels suitable for built public services may also contribute to the

wider network of public spaces, the analysis reveals a significant potential to support the planning of an interconnected system of public spaces and facilities, enhancing accessibility and addressing the current fragmentation and insufficient provision of public amenities across the study area.

As discussed in Sect. 4, suburban areas are particularly relevant for transformation, as they retain significant room for public intervention due to lower levels of land saturation and service provision, especially within the framework of the 15-min city. An illustrative application is provided by the suburban neighbourhood of Turano, a former satellite of the historic city, characterised by a mix of dispersed settlement patterns and a small, dense core currently lacking adequate collective space.

Considering only the top 20% of suitability values, the Aurelia Road clearly acts as a separator between two distinct urban systems. In the upper part, the compact historic fabric, together with the constraints associated with the foothill condition, primarily supports the allocation of open public spaces. In contrast, the lower part, characterised by lower land occupation and relatively larger parcels, shows a higher potential for the introduction of new public facilities, particularly in the area adjacent to the former wholesale market complex (Fig. 3). This spatial pattern reflects the combined effect of morphological, regulatory, and accessibility conditions captured by the model.

Even from a preliminary observation, a significant deficit of urban green spaces and collective gathering places can be identified, highlighting areas where local service provision could be improved. Notably, the areas adjacent to the school are currently occupied by parking lots, further emphasising the lack of accessible active public space. Considering the full set of most suitable areas, Fig. 3 highlights a high potential for public transformation, particularly in the area south of the Aurelia Road. Here, the concentration of high suitability values suggests the possibility of parcel aggregation, enabling the development of public spaces or facilities at a larger scale. At the same time, multiple opportunities emerge for the creation of an interconnected spatial network.

The following section presents a model representing one possible scenario for network development in this area, derived directly from the suitability-based approach outlined above.

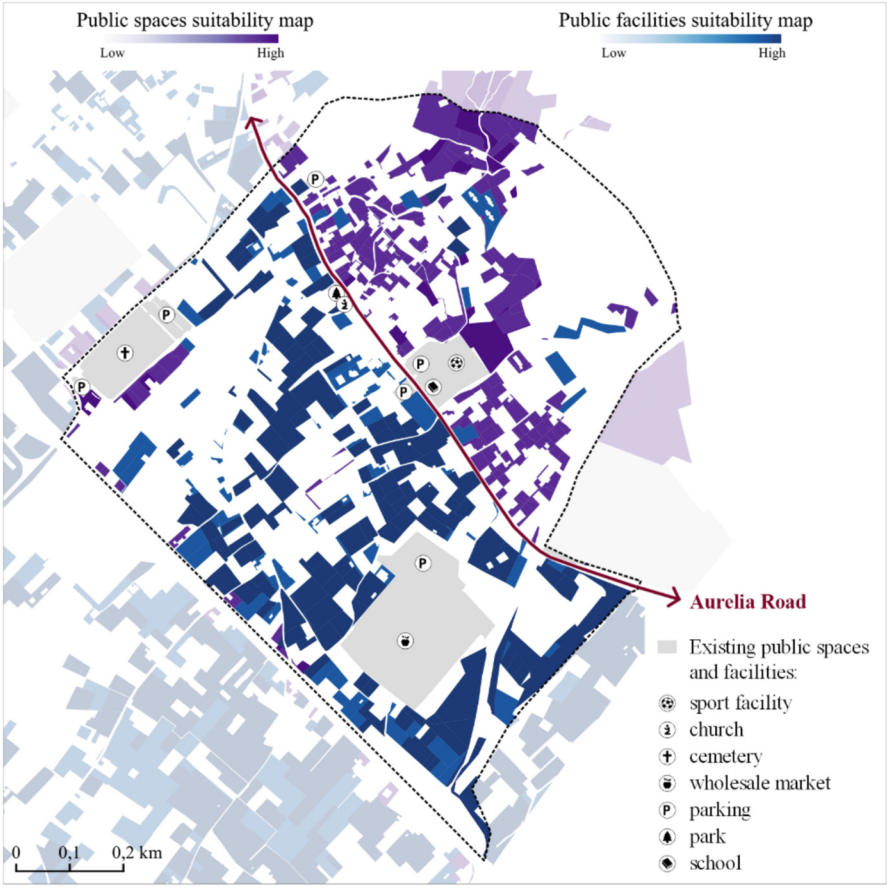


Fig. 3. Top 20% most suitable lots for public spaces and public facilities in the Turano study area

5.1 From Parcel Suitability to Facility–facility Proximity

This subsection visually demonstrates how parcel-level feasibility can be operationalised through a network model, aligned with the X-minute city framework, which treats the parcel as the generative unit and the network as the integrative mechanism.

Figure 4 provides an initial representation of the proposed facility–facility graph, in which facilities shown in Fig. 3 are modelled as nodes and any pair of facilities is linked by an edge where at least one walking route of less than 15 min exists between them. Outside the study area, connections to the railway station (to the north) and the heart hospital (to the south) have been included, due to the important role of these two infrastructures: the former supports the decarbonisation objectives underpinning the X-minute city model, while the latter has the potential to activate processes of attraction and revitalisation within the surrounding urban fabric.

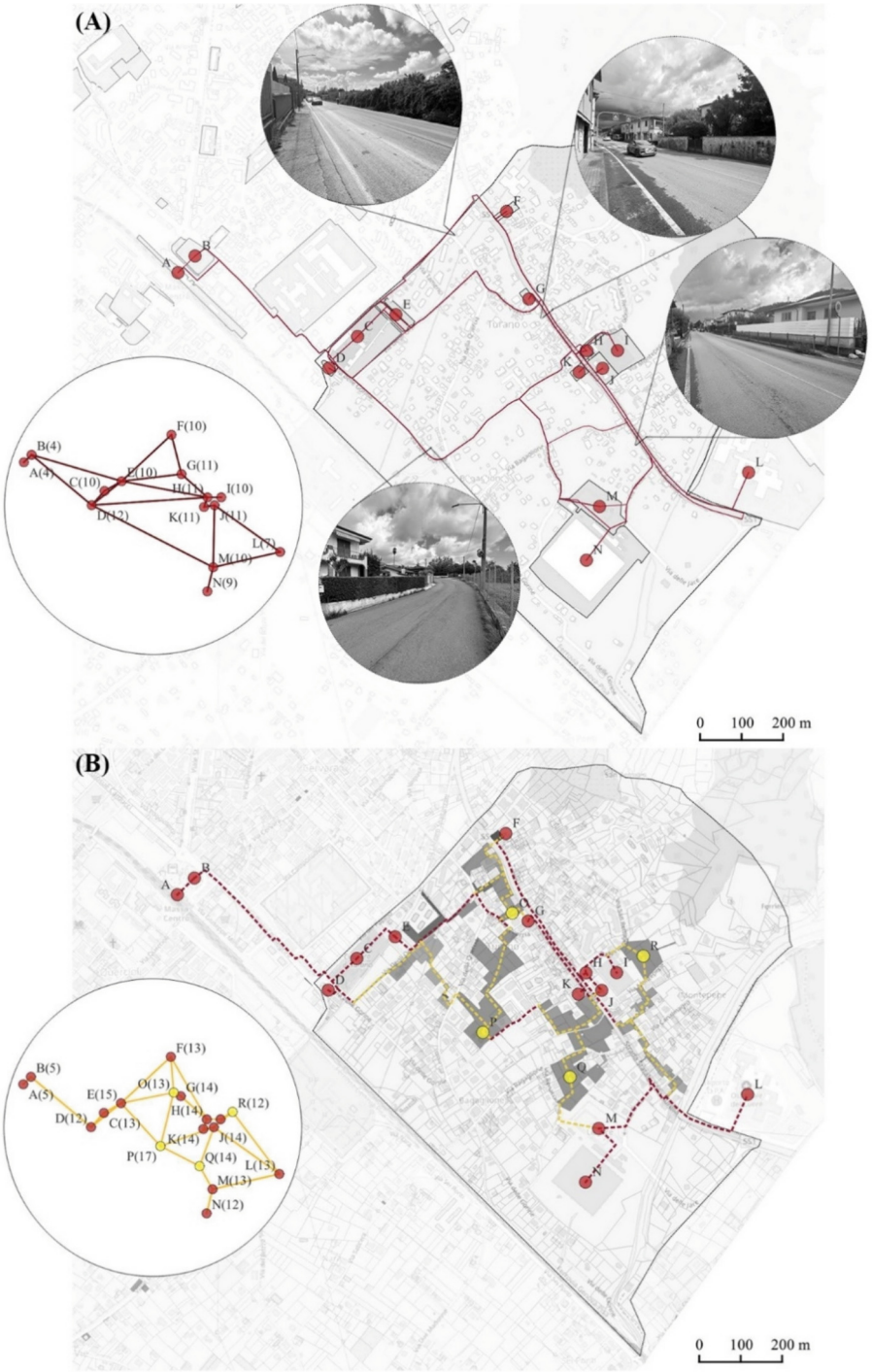


Fig. 4. Facility-facility network-based graph model. (A) Current situation; (B) Projected scenario to be assessed.

Figure 4a illustrates the current situation. Although the number of public facilities in Turano is limited, they can nonetheless be interpreted as a proximity network that establishes a baseline "scaffold" for subsequent, decision-oriented scenario-based evaluations. The map shows existing public facilities (red nodes) and the shortest walking connections under 15 min between them (red paths), as returned by *Google Maps* at the time of mapping (March 2026). In its current state, public spaces are unevenly distributed and largely dependent on a limited number of linear connections, with several terminal nodes. Moreover, while some connections fall within a 15-min threshold according to the street network, these routes are not necessarily walkable in practice. Field inspections (see the images in Fig. 4a) revealed a lack of basic pedestrian infrastructure: sidewalks are present along Via Aurelia but remain discontinuous, and are entirely absent on other roads.

Figure 4b then shows a candidate future scenario - to be assessed (e.g., in terms of coverage, cost, and facility centrality). Starting from the top 20% of the most suitable parcels, the scenario combines anchor nodes (existing facilities) with candidate nodes for urban space transformation, together with their network-based connections meeting the same travel-time threshold adopted in Fig. 4a. Connections are enhanced through the introduction of dedicated pedestrian- and cyclist-safe routes, linking both new and existing nodes. In both cases, the basemap is included only to orient the reader; no further spatial statistics, weights, or thresholds are applied beyond the 15-minute cut-off. This representation of the anchor layer (existing facilities), the edge layer (sub-15-minute walking ties), and scenario readiness (candidate facilities plus sub-15-minute walking links) provides a simple entry point for integrating the parcel-level suitability model with a network-based proximity reading, conveying:

- a skeleton of proximity, where existing services already form chains of sub-15-minute connectivity;
- connectivity gaps, where spatial separations exceed walkable thresholds;
- opportunity loci, where the addition of a node or the improvement of a link could bridge weakly connected clusters and strengthen facility-to-facility relationships.

Future work should, therefore, build on this qualitative baseline to support more quantitative evaluations (e.g., via network analytics in Python) of where dispersed fabrics and discontinuous public space would benefit most from targeted service nodes and last-mile pedestrian upgrades.

As noted, this study proposes a preliminary facility–facility graph. The numbers in brackets indicate the proximity-connectivity value of each node, calculated manually by counting the number of shortest paths linking that node to others within the 15-min threshold. In Scenario A, shortest paths are derived from *Google Maps*, whereas in Scenario B connectivity is estimated using path distances and an average walking speed of 4.5 km/h, commonly associated with adult pedestrians [34]. The comparison between the two graphs shows a substantial increase in connectivity in Scenario B, driven not only by the introduction of new public hubs but also by the addition of central connecting routes. Overall, the intervention increases network density, reduces structural vulnerability, and promotes a more cohesive urban system.

In this scenario, the lower area adjacent to the railway is slightly deprioritised, despite the high suitability of some parcels. This reflects emc2-model principles which favour

the concentration of services, activities, and urban life along the main street – here represented by Via Aurelia [8]. Nevertheless, routes that in Scenario A relied on roads within this area can be rerouted onto the new pedestrian-safe network without compromising their 15-min connectivity value.

6 Conclusion and Further Developments

This work proposes a parcel-based computational decision-support methodology for mapping urban spatial potential, with particular application to 15-min city planning. Rather than advancing an abstract or purely normative model of proximity, the framework seeks to identify realistic transformation scenarios by taking cadastral parcels as the operative unit and by assessing their effective predisposition to host public spaces and public facilities under existing spatial, morphological, regulatory conditions. In doing so, the methodology brings the 15-min city concept closer to the realities of implementation-oriented local planning, where public action must engage with the material and legal structure of land rather than with idealized spatial configurations. In this sense, the relevance of land ownership is not a secondary aspect, but a confirmation that urban transformation in suburban contexts largely takes place within a fragmented fabric of private property, where the feasibility of intervention depends on the actual constraints and opportunities embedded in land tenure, land use, and development conditions. For this reason, the proposed model is especially suited to local implementation planning, where public objectives need to be aligned with the practical conditions under which transformation can occur.

At the same time, the framework proves to be not only reproducible, thanks to its explicit workflow and criteria structure, but also flexible, since its scoring logic, thresholds, and decision rules can be calibrated according to different planning aims and municipal priorities. This makes it possible to adapt the model to support alternative policy orientations, ranging from more economically driven strategies to more sustainability-oriented scenarios attentive to ecological, social, and public-space quality objectives. In this sense, the framework does not provide a single fixed solution, but a methodological basis through which different public administrations can construct and compare scenario-based strategies according to their own goals and constraints.

Future research should therefore focus on strengthening the robustness and applicability of the framework by improving the integration and spatial consistency of heterogeneous datasets, validating the adopted criteria through expert review and stakeholder interviews, and further extending the model toward more advanced decision-support applications through the introduction of differentiated weights or exclusion thresholds. In particular, parcel-level suitability can be integrated with quantitative network analyses, including facility-level centrality measures (e.g., *closeness*, *betweenness*), as well as street-based representations able to preserve configurational metrics derived from network structure. At the same time, the 15-min accessibility perspective could be broadened beyond walking alone to account for multimodal accessibility and public transport connections, especially those related to strategic nodes such as railway stations. In this way, the model can provide the basis for a more comprehensive and context-sensitive evaluation of how public spaces and facilities may be progressively organized into interconnected proximity systems.

Acknowledgments. This work was part of the emc2 project, funded by ANR (France), FFG (Austria), MUT (Italy) and Vinnova (Sweden) under the Driving Urban Transition Partnership, which has been funded by the European Commission.

References

1. Mela, A.: Urban public space between fragmentation, control and conflict. City Territory Architect. **1**(1), 15 (2014)
2. Barone, R.E., Giuffrè, T., Siniscalchi, S.M., Morgano, M.A., Tesoriere, G.: Architecture for parking management in smart cities. IET Intel. Transport Syst. **8**(5), 445–452 (2014)
3. Buckland, M., Pojani, D.: Green space accessibility in Europe: a comparative study of five major cities. Eur. Plan. Stud. **31**(1), 146–167 (2023)
4. Cutini, V., Mara, F.: Many shades of green: intrinsic and network properties of urban green areas. TeMA-J. Land Use Mob. Environ. **1**, 147–167 (2025)
5. Allam, Z., Moreno, C., Chabaud, D., Pratlong, F.: Proximity-based planning and the "15-minute city": A sustainable model for the city of the future. In: The Palgrave Handbook of Global Sustainability, pp. 1523–1542. Springer International Publishing, Cham (2023)
6. De Vidovich, L.: Socio-spatial transformations at the urban fringes of Rome: unfolding suburbanisms in Fiano Romano. Eur. Urban Reg. Stud. **29**(2), 238–254 (2022)
7. Bruno, M., Monteiro Melo, H.P., Campanelli, B., Loreto, V.: A universal framework for inclusive 15-minute cities. Nature Cities **1**(10), 633–641 (2024)
8. Fusco, G., Pont, M.B., Cutini, V., Psenner, A.: Guiding principles for the 15-minute city in peripheral areas: the emc2 model. Game Changer? Plan. Just Sustain. Urban Regions, 690–707 (2024)
9. Micelli, E.: Development rights markets to manage urban plans in Italy. Urban Stud. **39**(1), 141–154 (2002)
10. Marchigiani, E., Bonfantini, B.: Urban transition and the return of neighbourhood planning. Questioning the proximity syndrome and the 15-minute city. Sustainability **14**(9), 5468 (2022)
11. Batty, M.: Agents, cells, and cities: new representational models for simulating multiscale urban dynamics. Environ Plan A **37**(8), 1373–1394 (2005)
12. Portugali, J.: Complexity, Cognition and the City, pp. 113–138. Springer, Berlin (2022)
13. Cozzolino, S.: On the spontaneous beauty of cities: neither design nor chaos. Urban Des. Int. **27**(2), 43–52 (2022)
14. Cozzolino, S., Moroni, S.: Structural preconditions for adaptive urban areas: framework rules, several property and the range of possible actions. Cities **130**, 103978 (2022)
15. Qin, S., Wang, C., Yan, Y.: Identification of conflict and its evolution between land use and land suitability during urban expansion: a case of Guangzhou, China. Ecol. Front. **44**(6), 1306–1319 (2024)
16. Murgante, B., Valluzzi, R., Annunziata, A.: Developing a 15-minute city: evaluating urban quality using configurational analysis. The case study of Terni and Matera, Italy. Appl. Geography **162**, 103171 (2024)
17. Yamu, C., Garau, C.: The 15-min city: a configurational approach for understanding the spatial, economic, and cognitive context of walkability in Vienna. In: International Conference on Computational Science and Its Applications, pp. 387–404. Springer International Publishing, Cham (2022)
18. Tümtürk, O., Karakiewicz, J., de Haan, F.J.: Measuring the impact of plot types on physical change: a diachronic analysis of urban form evolution in New York, Melbourne and Barcelona. Cities **154**, 105380 (2024)

19. Akay, M., Çalışkan, O.: A parametric approach to plot-based urban design: a climate-responsive algorithmic control for the generation of urban block. Urban Des. Int. 1–22 (2025)
20. Cozzolino, S., Moroni, S.: Multiple agents and self-organisation in complex cities: the crucial role of several property. Land Use Policy **103**, 105297 (2021)
21. Dardanelli, G., Maltese, A.: On the accuracy of cadastral marks: Statistical analyses to assess the congruence among GNSS-based positioning and official maps. Remote Sens. **14**(16), 4086 (2022)
22. Cienciała, A., Sobolewska-Mikulska, K., Sobura, S.: Credibility of the cadastral data on land use and the methodology for their verification and update. Land Use Policy **102**, 105204 (2021)
23. Roić, M., Križanović, J., Pivac, D.: An approach to resolve inconsistencies of data in the cadastre. Land **10**(1), 70 (2021)
24. D'Amico, B., Pomponi, F.: A compactness measure of sustainable building forms. Royal Soc. Open Sci. **6**(6), 181265 (2019)
25. Marincu, C., Dan, D., Moga, L.: Investigating the influence of building shape and insulation thickness on energy efficiency of buildings. Energy Sustain. Dev. **79**, 101384 (2024)
26. Tong, D., Yuan, Y., Wang, X., Wu, L.: Spatially varying relationships between land ownership and land development at the urban fringe: a case study of Shenzhen, China. Cities (2020)
27. Orr, A., Stewart, J., Jackson, C., White, J.: Ownership diversity and fragmentation: a barrier to urban centre resilience. Environ. Plann. B Urban Anal. City Sci. **50**, 660–677 (2022)
28. Gehl, J.: Cities for People. Island press (2013)
29. Jerrett, M., Su, J., MacLeod, K., Hanning, C., Houston, D., Wolch, J.: Safe routes to play? pedestrian and bicyclist crashes near parks in Los Angeles. Environ. Res. **151**, 742–755 (2016)
30. Benítez-López, A., Alkemade, R., Verweij, P.: The impacts of roads and other infrastructure on mammal and bird populations: a meta-analysis. Biol. Cons. **143**, 1307–1316 (2010)
31. Fedorca, A., Fedorca, M., Ionescu, O., Jurj, R., Ionescu, G., Popa, M.: Sustainable landscape planning to mitigate wildlife–vehicle collisions. Land (2021)
32. Porta, S., Crucitti, P., Latora, V.: The network analysis of urban streets: a primal approach. Environ. Plann. B. Plann. Des. **33**(5), 705–725 (2006)
33. Mara, F., Anselmi, C., Deri, F.: Urban systems, when they conurbate: the diachronic-configurational approach to detect the shifting of centrality in Versilia. URBAN DESIGN International, 1–22 (2025)
34. Pezzica, C., Altafini, D., Mara, F., Chioni, C.: Travel-time in a grid: modelling movement dynamics in the "Minute City". In: Marucci, A., Zullo, F., Fiorini, L., Saganeiti, L. (eds.) Innovation in Urban and Regional Planning. INPUT 2023. Lecture Notes in Civil Engineering, vol. 467. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-54118-6_58

Distribution of Spatial Inequalities Across Social Groups: a Case Study of Messina, Italy

Beniamino Murgante and Alfonso Annunziata(✉)

Department of Engineering, University of Basilicata, 85100 Potenza, Italy
alfonso.annunziata@unibas.it

Abstract. In response to contemporary urban crises, including social inequality, environmental degradation, and climate change, the 15-min city has emerged as a promising framework for fostering equity, social cohesion, resilience, and sustainability. Rooted in chrono-urbanism, this model seeks to improve quality of life by reducing travel time to essential urban functions. However, existing studies have largely focused on spatial accessibility, often overlooking broader dimensions of social justice such as digital inclusion, recognitional justice, and procedural participation. Addressing these gaps, this study adopts Fainstein's reform-oriented perspective on social justice and examines the extent to which existing cities comply with the principles of diversity, distributional equity, digital inclusion, and participation. To this end, it develops a quantitative and replicable methodology combining spatial and configurational analysis to assess access to eight categories of fundamental services, political participation, diversity, socio-economic disadvantage, and digital inclusion across six population groups differentiated by age and nationality. The city of Messina, Italy, is used as a case study. The main contribution of the study lies in the development of a composite social justice metric and in the proposal of a transferable evaluation framework capable of supporting transparent, evidence-based governance in the implementation of the 15-min city.

Keywords: 15-min city · social justice · spatial equity · participation · diversity

1 Introduction

Emerging crises affecting contemporary cities, including social inequality, environmental degradation, and climate change, have generated the need for a spatial reorganization of urban systems. Within this context, the 15-min city (15MC) concept has emerged as a relevant model promoting equity, social cohesion, resilience, and sustainability in urban areas. Grounded in the concept of chrono-urbanism, the 15-min city model posits that quality of life is closely related to patterns of time use and can be improved by reducing the time required to travel in order to access fundamental urban functions [1–5]. Consequently, the model envisions an organization of urban areas in which essential functions are located within a radius corresponding to a travel time of approximately 15 min on foot or by bicycle from one's residence [4, 5]. In this perspective, the conceptualization of the 15-min city incorporates principles associated with social justice: equity, understood as the fair distribution of resources; diversity, referring to the co-presence and

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 22–38, 2027.
https://doi.org/10.1007/978-3-032-30530-5_2

inclusion of different identities, cultures, and social groups; and participation, intended as the transparency and inclusiveness of governance and political processes [6].

Nevertheless, emerging critiques of the 15-min city concept observe that only a limited number of studies explicitly investigate the justice implications of the model, and that even those addressing the topic often do not ground their analyses in established conceptualizations of social justice. Moreover, additional critiques underline the insufficient consideration of the social, economic, and political factors determining injustice, as well as the predominant focus on spatial inequalities related to service distribution and accessibility. This emphasis tends to disregard complementary dimensions of justice, including exclusion resulting from the digital divide, recognitional justice, and procedural justice [6]. To address these limitations, the study builds on the liberal, reform-oriented perspective on social justice advanced by Fainstein [7] – based on a conceptualization of social justice around the dimensions of equity, democracy, and diversity - and investigates a specific question: to what extent existing cities are compliant with principles of diversity, distributional equity, digital inclusion and participation?

To answer this question, the proposed study develops a quantitative method that utilizes spatial analysis and configurational analysis to evaluate levels of access to eight categories of fundamental services, political participation levels, diversity, socio-economic disadvantage, and digital inclusion, across six distinct population groups defined by age and nationality. The city of Messina, in Italy, is selected as a case study. The relevance and novelty of the contribution lie in several distinct aspects. First, it contributes to the operationalisation of social justice, as conceptualised by Fainstein, by developing a composite spatial metric that integrates the dimensions of participation, digital divide, diversity, socio-economic disadvantage, and spatial equity.

Second, it proposes a replicable, robust, and transferable procedure for evaluating conditions of spatial justice across urban areas and urban populations. The objective is to foster transparent and evidence-based governance processes in the implementation of the 15-min city model - in order to mitigate the risk of reproducing forms of marginalisation and inequality - by enabling the understanding of conditions of intra-urban disparities. The article is structured into six sections. After the introduction, Sect. 2 elaborates on the conceptual foundations, while Sect. 3 details the proposed methodology. Section 4 provides an overview of the results, which are subsequently examined in Sect. 5. The Conclusions summarize the principal findings and offer suggestions for future research directions.

2 Literature Review

2.1 Definitions of Social Justice

Ensuring justice in decisions related to spatial transformations represents a central concern of urban policies [8]. Justice is not only a fundamental policy objective in itself, but the claim of pursuing an equitable distribution of benefits and disadvantages also confers legitimacy on policy decisions. In this context, the notion of social justice has emerged as a central analytical and normative concept for evaluating urban policies and spatial planning processes. Social justice focuses on the effect of social structures in reproducing or mitigating inequalities across the dimensions of resource distribution, inclusivity

of political processes, and recognition of diversity [6]. The definition of social justice has evolved, over time, from a predominant focus on distributional equity to broader conceptualizations integrating aspects of procedural and recognitional justice. The recent literature identifies multiple definitions of social justice, including utilitarian, liberal egalitarian, Marxist, and libertarian perspectives.

A Marxist perspective is intrinsic to definitions of social justice advanced by Marcuse and Lefebvre. Marcuse [9] identifies specific spatial aspects of injustice – namely confinement and unequal distribution of resources – and conceptualizes spatial injustice as a derivative of social injustice determined by specific social, political, and economic conditions. Lefebvre's notion of the right to the city [10] underlines individuals' entitlement to participate in decisions structuring urban space and to access and produce spaces responding to social needs. Similarly, Harvey [11] argues for the incorporation of social justice considerations into procedures addressing locational problems, proposing the operationalization of concepts such as need, contribution to the general good, and merit as criteria for favouring just distributions of spatial resources.

A critical liberal egalitarian perspective is advanced by Rawls and developed in urban studies by Soja and Fainstein. Rawls conceptualizes justice in terms of the equitable distribution of resources and goods. Building on this perspective, Soja focuses on the spatial dimension of justice/ injustice and conceptualizes spatial injustice both as an outcome - manifested in uneven distributional configurations - and as the set of processes that produce injustice via the organization and control of space [12–14].

Fainstein's concept of the Just City [7] contributes to the operationalization of spatial justice by proposing justice as the normative principle guiding urban policies. In this context, democracy, equity, and diversity constitute the fundamental principles of urban justice [7].

The present study adopts a liberal perspective, which underlies many empirical studies investigating spatial inequalities in the distribution of urban amenities and infrastructures, namely transport systems [15], Services of General Interest [16–18], and natural and semi-natural spaces[19]. Within this perspective, the question of spatial justice is intrinsic to recent urban concept, particularly the 15MC and the x-minute City, as observed by Murgante et al. [16–18].

2.2 The 15-Minute City Concept

The concept of the 15-min city (15MC) - more broadly framed as the x-minute city [20]- was introduced by Moreno in 2016 [2, 5, 21] and has been adopted by municipal governments and transnational initiatives as a model for adapting cities to emerging crises including climate alteration, environmental degradation and social injustice. Examples include initiatives promoted by the C40 Cities initiative and the implementation of related policies in several cities, including Milan, Paris, Bogotá, Portland, and Melbourne [22, 23].The 15MC concept refers to a model of socio-spatial reorganization of urban areas aimed at ensuring fair access to essential services and urban amenities by promoting active mobility and reducing travel times. The model aims to reorganize urban functions so that fundamental needs can be met within a distance from one's residence corresponding to a travel time of 15 min on foot or by bicycle.

Expected benefits include improvements in quality of life, social inclusion, economic vitality, ecological resilience, and spatial justice [2, 16–18, 24–26].

Moreno identifies four central criteria for the implementation of the 15MC: proximity, density, diversity, and digitalization. Proximity refers to the optimal spatial distribution of urban functions in order to facilitate access to services, while promoting active mobility and reducing dependence on private cars. Diversity concerns the coexistence and integration of varied activities, functions, and social identities at the local scale. Density refers to the optimal number of people per unit area that can be sustained in terms of service provision and resource consumption [5, 25]. Finally, digitalization involves the use of Information and Communication Technologies (ICT), Internet of Things (IoT) systems, distributed sensors, and digital twin platforms to optimize service provision, increase transparency and participation in governance processes, promote sustainable mobility solutions, respond to the needs of vulnerable users, and improve the safety, comfort, and usability of public spaces [16–18, 27]. Additional criteria discussed in the literature include trans-scalarity, intermodality, adaptability, and human-scale urban design [2, 21]. Trans-scalarity refers to the interdependence among different spatial levels - micro, meso, and city scale. Human-scale urban design focuses on the regeneration of public spaces centered on person–environment interactions, and proportion of urban space elements in relation to the human body and its capabilities. Intermodality refers to the integration of public transport systems and active mobility solutions. Criteria of decentralization and diversity of urban functions, proximity of Services of General Interest, optimization of density, access to natural and semi-natural spaces, compact urban structure, and the prioritization of active mobility and pedestrian-oriented development are also central to several earlier urban models. These include the Neighbourhood Unit Concept (NUC) [28], Transit-Oriented Development (TOD) [29, 30], New Urbanism [31], neo-traditionalist design concepts [32], multi-fractal urban models [33], and the conceptualizations of urbanity and urban vitality advanced by Jane Jacobs [34], Christopher Alexander [35], and Jan Gehl [36].

Considering these antecedents, the 15-min city does not represent an entirely new paradigm. Rather, it can be interpreted as a contemporary operational reframing of consolidated principles and longstanding normative aspirations concerning proximity, mixed-use development, and human-scale urbanism. Moreover, the 15-min city model has been subject to significant criticism. The most prominent critiques concern the conceptual framing of the model, inconsistencies in the quantitative methods employed to operationalize and measure its foundational dimensions, and questions regarding its practical feasibility [37–39]. Lastly, critical perspectives underline the inadequate conceptualization and operationalization of social justice aspects and the risk that the implementation of the model may reproduce or exacerbate existing socio-spatial inequalities [37, 40].

3 Methodology

The study advances a quantitative method that combines spatial and configurational analysis to compute a composite metric capturing factors of social injustice and to assess levels of distributional equity, digital inclusion, participation, socio-economic status,

and diversity across six distinct population groups. The procedure is organised into four stages: (i) data collection; (ii) selection and computation of indicators; (iii) computation of a composite spatial justice metric; and (iv) estimation of the distribution of adequate spatial justice conditions across urban population groups. The procedure is applied to the urban area of Messina (Fig. 1), which provides an appropriate empirical setting for examining socio-spatial inequality through the interpretive lens of the 15-min city. Messina is a medium-sized city located at the north-eastern tip of Sicily. Covering approximately 213.75 km^2 and with a population of 216,566 inhabitants, the city combines a relatively dispersed municipal territory with a dense urban core structured around the port, the waterfront, and the main axes of connection across the Strait.

Its contemporary spatial configuration is strongly influenced by the post-1908 reconstruction, which introduced a rationalized street layout and new building typologies, partly replacing the pre-existing historical fabric. Subsequent planning phases, including the 1976 Tekne plan, the 2002 master plan, and the current updating of the Urban Plan (PRG), have further contributed to uneven patterns of urban growth, land consumption, and access to services. These characteristics make Messina a highly relevant case study for investigating conditions of social injustice through the interpretive lens of the 15-min city. Specifically, the city's complex relationship between topography, coastal infrastructure, neighbourhood distribution, public space, and service provision offers a useful setting for assessing whether everyday urban opportunities - such as accessibility to essential services, mobility options, and collective facilities - are equitably distributed across different populations and areas. In this perspective, Messina can be understood as a critical urban laboratory in which historical legacies, spatial fragmentation, and contemporary planning challenges intersect, making visible the territorial dimensions of socio-spatial inequality.

3.1 Data Collection

Data are retrieved from open data platforms to ensure the replicability and reproducibility of the analysis. Population data at the census tract level - including the total number of residents and the number of residents by age, gender, education level, employment status, and nationality - are obtained from the 2021 Census provided by the Italian National Institute for Statistics [41]. Concerning Services of General Interest, the categories considered include: grocery stores; pharmacies; secondary schools; public transport stops; green areas; retail activities (including bookshops, clothing stores, and personal care retail); general services (including post offices and banks); and food and beverage service activities. Points of Interest (POI) data are collected from Google Maps using the G Maps Extractor browser extension (version 2.4.7).

Google Maps is adopted as the primary data source for Services of General Interest (SGIs) due to its extensive spatial coverage, enhancing the transferability and replicability of the proposed procedure. Nevertheless, data derived from Google Maps may present limitations related to accuracy, completeness, and classification consistency [42]. Data on educational services are obtained from the Italian Ministry of Education platform. Data are georeferenced and verified for positional accuracy. Spatial data on green areas are retrieved from the OpenStreetMap platform and from the land use–land cover layer of the Urban Atlas service of the Copernicus platform. Points representing access to urban

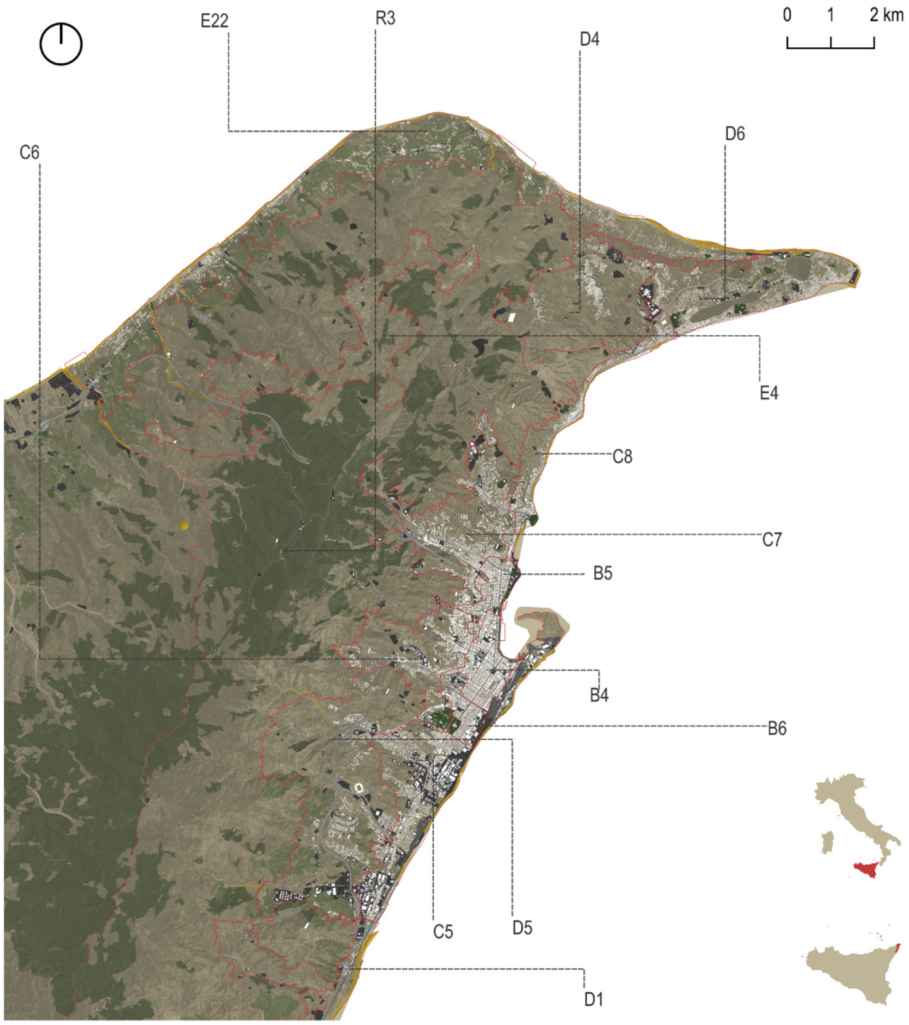


Fig. 1. Description of the study area: the urban area of Messina.

green areas are generated by intersecting the perimeters of green spaces with segments representing pedestrian and cycling routes; resulting points are retained as destination locations. The urban area is partitioned into 16 zones according to the classification of the Real Estate Market Observatory (OMI) of the Italian Revenue Agency. These zones are categorised as: B - central areas; C - semi-central areas; D - suburban areas; E - peri-urban areas; and R - rural areas. This classification is adopted here as a spatial proxy to capture intra-urban variation in spatial and socio-economic conditions.

The system of urban roads, and streets is represented as a segment map derived from a road centreline dataset retrieved from the OpenStreetMap platform. The road centreline layer is verified for topological errors and edited by removing geometries

representing spaces unsuitable for pedestrian or cycling mobility. In particular, segments labeled as motorway, trunk, service roads, and tracks are excluded. The remaining lines are converted into segments using the Place Syntax Tool by partitioning geometries at intersections. The resulting segment map is then verified to identify and remove isolated segments. Regarding broadband service availability, data on expected download speed at the 20 m grid cell level are retrieved from the dataset provided by the Italian Communications Regulatory Authority (AGCOM) [43]. For civic participation, voter turnout data are obtained from the open data platform of the Municipality of Messina. The number of registered voters and the number of actual voters in the most recent election are aggregated at the electoral section level. Voter turnout is used as an indicator of local levels of civic participation.

Lastly, data on socio-economic disadvantage are retrieved from the dataset provided by the Italian National Institute of Statistics (ISTAT), which identifies census tracts categorized as areas of socio-economic vulnerability. These areas are defined by aggregating contiguous census tracts with similar socio-economic conditions, measured via the IDISE metric - a composite indicator based on nine variables representing fundamental dimensions of socio-economic disadvantage. Socio-economic disadvantage is defined as a condition of inadequate access to social, economic, employment, and educational resources that prevents individuals from meeting their basic needs.

3.2 Selection and Computation of Indicators

Five dimensions of spatial justice - namely distributional equity, diversity, socio-economic status, civic participation, and digital inclusion - are operationalised in terms of five binary variables computed at the census tract level. Each variable assumes a value of 1 when a predefined condition of justice is satisfied and 0 when a condition of injustice is observed.

For the dimension of distributional equity, the minimum requirement is that at least two destinations for each category of Services of General Interest are located within a distance of 800 m from each census tract. The minimum value of two accessible destinations is adopted to account for a basic level of choice and redundancy in service availability. The number of accessible destinations within the predefined radius of 800 m is computed using the Attraction Reach function of the Place Syntax Tool, which measures the number of points of interest located within a specified network distance along the segment map. The 800m threshold is adopted to measure access to Services of General Interest (SGIs) within the 15-min city perspective, as 800 m approximately correspond to a travel time of 15 min at a speed of 3 km/h. This value is used as a conservative estimate of pedestrian movement across different population groups. For each service category, a destination attractiveness index is computed for every point of interest by combining three variables: (i) service quality, measured as the product of the number of user reviews and the average user rating; (ii) angular integration at a radius of 1000 m of the closest street segment; (iii) distance from infrastructural barriers.

These variables are selected to represent the intrinsic quality of destinations and the ease of pedestrian access to each point of interest. In particular, local angular integration is used as a metric of pedestrian accessibility, as it measures the relative centrality of street segments within a spatial system and has been positively associated with levels of

pedestrian movement and spatial accessibility in space syntax studies [44, 45]. Angular Integration is computed at a 1000 m radius to represent configurational properties at the neighbourhood scale, whereas the 800 m threshold adopted in the Place Syntax analysis captures local conditions of access to services corresponding to a 15-min travel time. The use of different radii reflects the distinct analytical purposes of the two measures.

The variables are normalised using range standardisation (min–max normalisation) and aggregated via arithmetic averaging to obtain the composite attractiveness index. Only destinations with index values above the mean are retained to ensure that access is evaluated with respect to destinations of adequate and comparable attractiveness. Civic participation is operationalised in terms of voter turnout in the most recent election. The condition of civic participation is considered satisfied when voter turnout in the closest electoral section is equal to or greater than 50% of registered voters. This value identifies areas where a minimum level of electoral engagement is observed. The digital divide dimension is operationalised using broadband infrastructure availability. The condition is considered satisfied when an expected broadband download speed of at least 1000 Mbps is available across at least 80% of the surface area of each census tract. This criterion ensures that advanced digital infrastructure is accessible to the majority of the tract area. Socio-economic disadvantage is evaluated using the IDISE index provided by the Italian National Institute of Statistics (ISTAT). The condition is considered satisfied when a census tract is not included within areas classified by ISTAT as zones of socio-economic disadvantage. Lastly, social diversity is measured through a multidimensional index computed at the census tract level. For each tract, Simpson's diversity index is calculated separately for age structure, educational level, employment status, and nationality across the focal tract and contiguous census tracts. The four dimension-specific scores are then normalised and aggregated through arithmetic averaging to obtain a composite indicator of multidimensional social diversity. The multidimensional social diversity index is subsequently converted into a binary variable, assuming that a condition of adequate social diversity is satisfied when the index value exceeds the median. Lastly, a composite metric is computed as the sum of the five variables, capturing the simultaneous satisfaction of the conditions of justice. A value of 5 indicates that all conditions are satisfied, whereas a value of 0 denotes the co-presence of five dimensions of social injustice.

3.3 Spatial Justice Across Populations

The final stage of the analysis evaluates the distribution of spatial justice conditions across different population groups. The resident population is classified into six groups defined by nationality and age: Italian residents aged 0–14 years, Italian residents aged 15–64 years, Italian residents aged 65 years and over, foreign residents aged 0–14 years, foreign residents aged 15–64 years, and foreign residents aged 65 years and over.

For each group, the analysis calculates the percentage of individuals residing in census tracts that exhibit adequate conditions of spatial justice across five dimensions: distributional equity, socio-economic status, digital inclusion, civic participation, and diversity. The results of this assessment are presented in the sub-sequent section.

4 Results

The spatial distribution of justice conditions indicates that census tracts presenting adequate conditions across the five selected dimensions are primarily located in the central compact core of the Messina urban area (Fig. 2), comprising zones B4, B5, B6 and the semi-central area C6. However, some census tracts located in central areas - particularly in zone B4 - also present conditions of marginalisation and segregation, in particular resulting from inadequate levels of diversity and civic participation. Census tracts presenting adequate conditions of socio-spatial justice across four out of the five dimensions are mainly situated at the interface between central and semi-central areas, particularly in central zone B4, and along the coastline in the semi-central area C8 and in the suburban zone D6.

Moreover, isolated clusters of census tracts characterized by mild conditions of injustice are located in suburban and peri-urban areas, including zones D1, E1 and E2.

Areas presenting conditions of social injustice ranging from severe to intense (indicator values equal to 1 and 2) are distributed across semi-central, peri-urban and suburban areas. Notably, multiple dimensions of deprivation – related to inadequate service provision, socio-economic status, and limited civic participation - are observed along the interface between central and semi-central areas, particularly between zones B6 and C5.

The distribution of justice conditions across population groups reveals substantial levels of deprivation, as well as distinct configurations of disadvantage affecting specific groups (Fig. 3). In particular, only 16.9% of the resident population can access at least two alternative destinations for each category of service. Similarly, on average only 38% of residents live in diverse census tracts, indicating relatively limited socio-spatial diversity across the study area.

More favorable conditions emerge in other dimensions. With respect to socio-economic status, 92% of individuals reside in areas classified as not vulnerable to socio-economic disadvantage. Positive outcomes are also observed for digital inclusion, with 85.5% of individuals living in areas where at least 80% of the surface is covered by broadband services with a download speed of 1,000 Mbps. Civic participation also presents relatively positive conditions, as 87.6% of individuals reside in census tracts where voter turnout exceeds 50%.

However, the analysis underscores relevant differences across population groups (Table 1). In relation to service provision, the most deprived group is represented by Italian residents aged 0–14 years, whereas foreign residents in the same age group appear to be the least disadvantaged. More precisely, 14.1% of Italian residents aged 0–14 and 35.8% of foreign residents in the same age group reside in areas with adequate levels of access to Services of General Interest.

Regarding socio-economic status, the proportion of individuals living in areas categorized as not disadvantaged varies across groups, ranging from 85.1% among foreign residents aged 65 years and over to 93% among Italian residents in the same age group. Substantial inter-group disparities are also observed in relation to civic participation: 88.4% of Italian residents aged 0–14 years reside in census tracts with voter turnout above 50%, compared to 78.9% of foreign residents aged 65 years and over.

Conditions of digital inclusion also vary across groups. Italian residents aged 15–64 years represent the most disadvantaged group in this dimension, whereas foreign

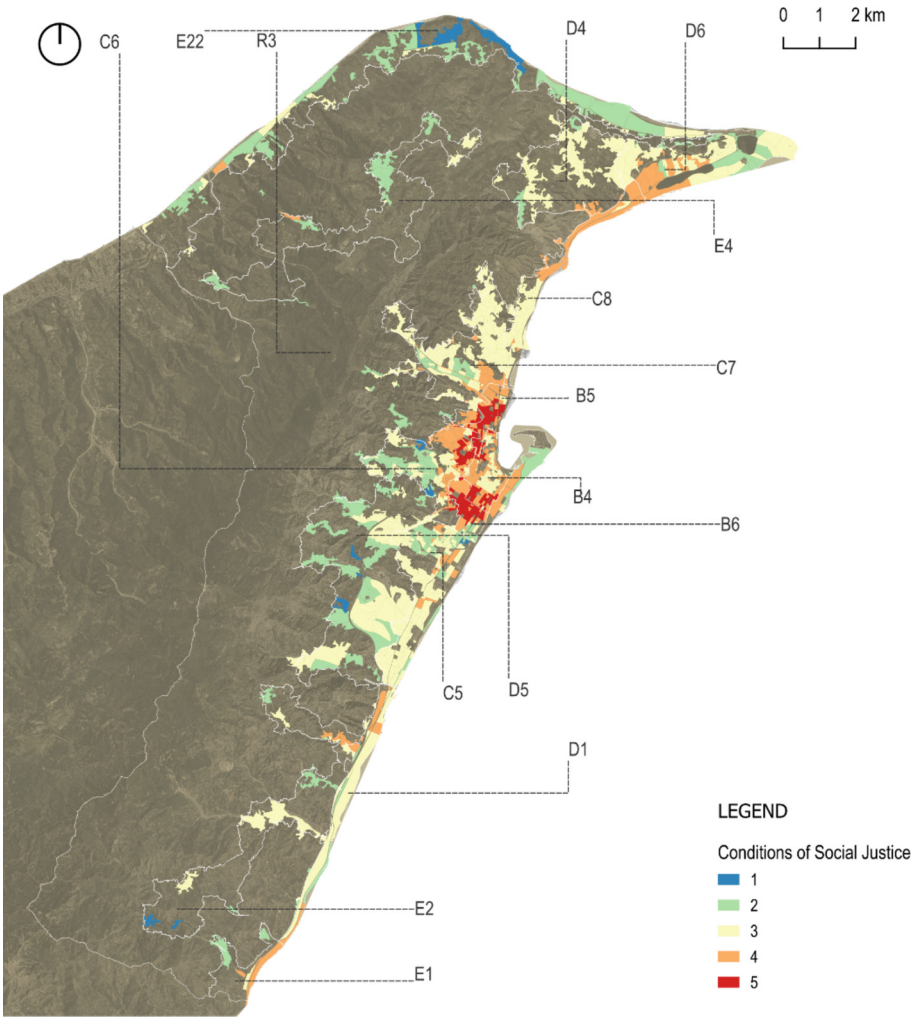


Fig. 2. Co-presence of conditions of social justice across census tracts in the city of Messina

residents aged 0–14 years appear to be the least disadvantaged, with respectively 85.1% and 92.4% of individuals residing in areas with adequate access to broadband infrastructure. Lastly, notable disparities emerge with regard to socio-spatial diversity. Adequate conditions of diversity are experienced by 35% of Italian residents aged 0–14 years, compared to 77.6% of foreign residents in the same age group. These results provide evidence on the mechanisms informing the uneven distribution of social justice conditions across population groups, as discussed in the subsequent section.

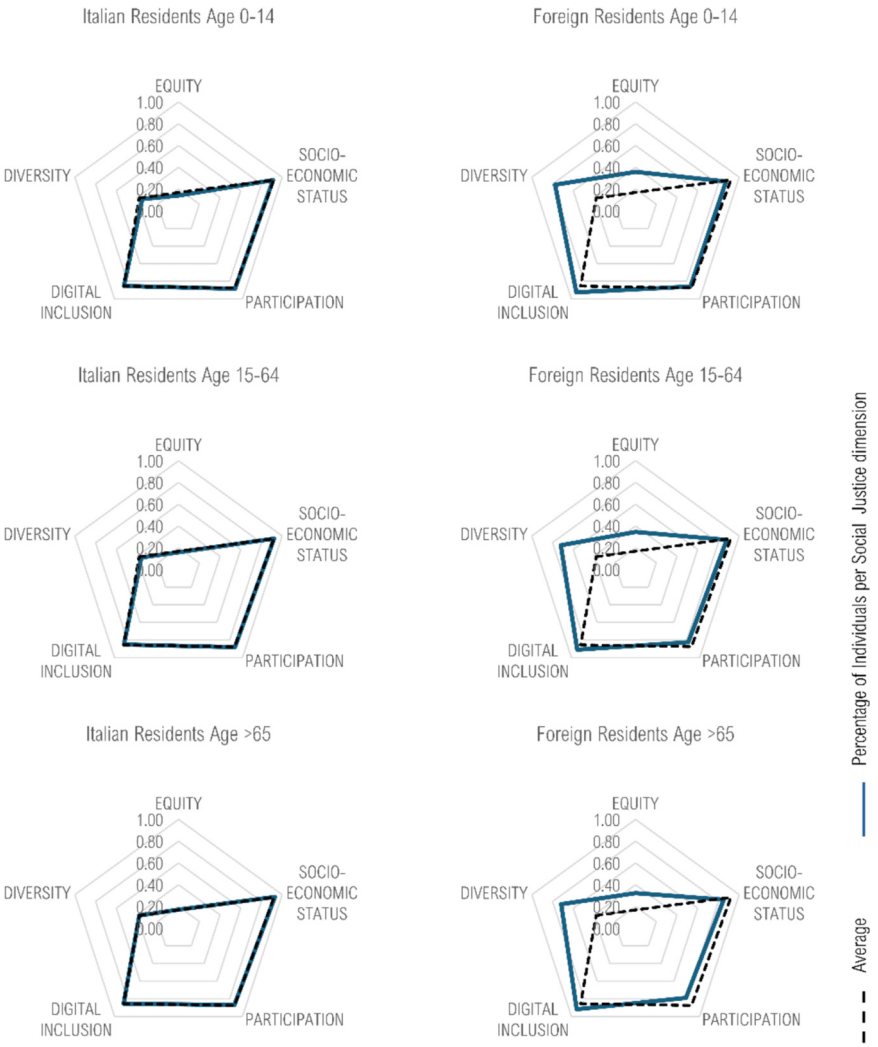


Fig. 3. Distribution of conditions of social justice across Population groups

Table 1. Distribution of conditions of social justice across six groups: Italian residents of age 0–14 years (IT_0_14), Italian residents of age 15–64 years (IT_15_64), Italian residents of age of 65 years and above (IT_OV_65), Foreign residents of age 0–14 years (FR_0_14), Foreign residents of age 15–64 years (FR_15_64), and Foreign residents of age of 65 years and above (FR_OV_65).

	Distributional Equity	Socio-Economic Status	Civic Participation	Digital Inclusion	Diversity
IT_0_14	0.141	0.912	0.884	0.852	0.350
IT_15_64	0.159	0.921	0.881	0.851	0.358
IT_OV_65	0.172	0.930	0.871	0.856	0.377
FR_0_14	0.358	0.874	0.859	0.924	0.776
FR_15_64	0.345	0.880	0.825	0.910	0.720
FR_OV_65	0.323	0.851	0.789	0.919	0.717
AVERAGE	0.169	0.920	0.876	0.855	0.380

5 Discussion

The results underscore several relevant patterns in the spatial and social distribution of urban justice across the study area. First, concerning the spatial distribution of justice conditions, the analysis reveals contrasting trends. On the one hand, a clear intra-urban divide emerges between central areas and suburban and peri-urban areas. On the other hand, central areas display internal heterogeneity, resulting in the co-presence of clusters of census tracts presenting multiple forms of deprivation and segregation alongside areas with favorable conditions. At the same time, clusters of census tracts presenting mild conditions of injustice are observed in suburban and peri-urban areas.

As a consequence, the analysis reveals a general trend resulting in a monocentric configuration. This configuration appears to be largely driven by the concentration of Services of General Interest in central areas and by elevated levels of access to advanced broadband infrastructure in denser urban zones. This trend reflects configurations observed in previous studies on the 15-min city [17, 18, 46]. However, the results also reveal substantial disparities within individual zones. In particular, intra-zone inequalities are mainly related to uneven conditions of social diversity and civic participation.

Regarding the distribution of social justice conditions across the urban population, the analysis identifies distributional equity and diversity as particularly critical dimensions. Access to urban resources and the presence of socially diverse environments remain limited and inadequate for a large segment of the urban population. Substantial disparities emerge across population groups, indicating uneven exposure to marginalisation and segregation. Italian residents are more prone to inequalities in the distribution of Services of General Interest and by spatial segregation. The percentage of Italian individuals residing in areas with adequate access to services remains limited across all age groups, ranging from 14.1% among residents aged 0–14 years to 17.2% among residents aged 65 years and above. By contrast, the corresponding values for foreign residents amount

to 35.8% among individuals aged 0–14 years and remain above 30% for the other age groups.

Foreign residents, however, face greater exposure to socio-economic disadvantage and lower levels of civic participation. More than 91% of Italian residents live in areas not exposed to conditions of socio-economic disadvantage, compared to percentages of foreign residents that range from 85.1% to 88.0%. A similar pattern is observed in relation to civic participation: the proportion of individuals residing in census tracts with voter turnout above 50% remains consistently larger among Italian residents than among foreign residents. Foreign residents also experience more favourable conditions in relation to digital inclusion and socio-spatial diversity. These patterns indicate that population groups are exposed to distinct dimensions of social justice. The observed disparities are closely related to the spatial distribution of urban populations. Foreign residents are more concentrated in central areas, where access to services, digital infrastructure, and diverse social environments is greater. Proximity to services and public transport facilitates access to fundamental urban resources, mitigating reliance on car-dependent mobility.

The Italian population, by contrast, is more spatially dispersed across the urban area. This pattern may reflect the combined influence of structural constraints and residential preferences. The presence of Italian residents in inland areas distant from the coastline, where service provision is more limited, may partly result from economic constraints that restrict residential options. In such contexts, lower housing costs may contribute to the concentration of lower-income residents. At the same time, suburban coastal areas may attract residents due to proximity to the seaside, lower-density environments, and alternative housing typologies. In these contexts, access to urban resources depends more strongly on private car use because of the limited availability of nearby services and public transport. This spatial configuration raises significant social justice concerns, including forms of segregation related to filtering-up processes in coastal areas, marginalisation of vulnerable groups in deprived suburban areas, unequal access to fundamental urban resources, and the persistence of dependence on private motorized transport. These findings call for an integrated, place-based policy package aimed at reducing the spatial concentration of advantages and disadvantages across the urban area by coupling land-use planning, service provision, housing instruments, and governance measures. In suburban and peri-urban areas, targeted densification - particularly in zones contiguous to existing service concentrations - can reinforce the demand base and urban intensity needed to support the decentralisation of Services of General Interest. This action should be complemented by the creation and consolidation of local centralities through the conversion of existing public amenities into multifunctional hubs that combine social, administrative, cultural, and civic functions, thus increasing proximity to urban resources [16, 27].

At the same time, segregation and limited socio-spatial diversity require housing policies that integrate social housing provision into development and regeneration processes. This objective can be advanced via public provision in targeted areas, mandatory or negotiated on-site social housing requirements in private redevelopment, and in-lieu contributions - including financial payments or land transfers - alongside density-based incentives that preserve project feasibility [47]. In combination, these instruments can

promote social mix in regeneration areas and expand access to affordable housing. Lastly, uneven civic participation indicates that spatial and distributive interventions should be supported by more inclusive governance structures [48, 49]. Participation in local governance and the co-design of regeneration interventions can foster civic engagement and the capacity of urban policies to recognise and respond to the needs of diverse social groups.

6 Conclusions

The study proposes a quantitative procedure based on spatial and configurational analysis to assess justice across five dimensions: distributional equity, digital inclusion, civic participation, diversity, and socio-economic status. The procedure is intended to support evidence-based and transparent governance in the design of policies oriented to the 15-min city model, with the specific aim of mitigating the reproduction of intra-urban disparities.

The contribution of the study is twofold. First, it advances the operationalization of social justice by translating Fainstein's definition [7] into a spatially explicit set of measurable dimensions. Second, it specifies a procedure that is replicable, robust, and transferable for assessing spatial justice conditions across urban areas and populations. The results provide empirical evidence that the procedure supports the systematic identification of spatial inequalities and the incorporation of justice considerations into policy design and evaluation. The findings indicate dimension-specific configurations of disparity across spatial units and population groups. More precisely, the dimensions of social justice diverge in terms of spatial gradients and inequality intensity, producing heterogeneous combinations of segregation and deprivation across the urban area and uneven exposure across population groups. In line with previous studies on the 15-min city [17, 18, 46, 49], distributional equity exhibits a monocentric configuration: Services of General Interest are concentrated in compact central areas, producing a centre-edge divide. Beyond this general pattern, the analysis also identifies intra-zone disparities across central, semi-central, suburban, and peri-urban areas. These intra-zone contrasts are most evident for civic participation and social diversity, where census tracts with favourable conditions coexist with tracts experiencing deprivation, marginalisation, or segregation.

Across population groups, the results indicate distinct profiles of exposure across the assessed dimensions. On this basis, the study advances the hypothesis that residential location - and the resulting exposure to specific dimensions of spatial injustice - derives from the interaction between the spatial distribution of the residential built environment and housing-price gradients and individual preferences and socio-economic constraints that delimit feasible residential options. This interpretation corroborates the intrinsic spatial nature of social justice, as articulated by Soja [12–14], insofar as residential location mediates exposure to uneven distributions of urban resources.

From a planning perspective, the results support an integrated, place-based policy orientation coordinating land-use regulation, service provision, housing instruments, and inclusive governance. The findings justify measures aimed at consolidating local centralities beyond the compact core, integrating social housing into development and

regeneration processes, and combining spatial interventions with participatory structures and co-design processes capable of addressing uneven civic participation.

Several limitations define priorities for future research. Further refinement is required to develop more pertinent indicators of the quality of Services of General Interest, more robust operationalizations of civic participation, and more fine-grained representations of socio-economic disadvantage. A second priority concerns empirical investigation of the mechanisms underlying the observed patterns, with specific attention to the relationships among socio-economic status, residential preferences, the spatial distribution of residential buildings and housing prices, and exposure to the assessed dimensions of social justice.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Allam, Z., Khavarian-Garmsir, A.R., Lassaube, U., Chabaud, D., Moreno, C.: Mapping the implementation practices of the 15-minute city. Smart Cities **7**, 2094–2109 (2024). https://doi.org/10.3390/smartcities7040083
2. Khavarian-Garmsir, A.R., Sharifi, A., Sadeghi, A.: The 15-minute city: urban planning and design efforts toward creating sustainable neighborhoods. Cities **132** (2023). https://doi.org/10.1016/j.cities.2022.104101
3. Biraghi, C.A., Ogut, O., Dong, T., Tadi, M.: CityTime: a novel model to redefine the 15-minute city globally through urban diversity and proximity. Urban Sci. **9** (2025). https://doi.org/10.3390/urbansci9020036
4. Allam, Z., Bibri, S.E., Chabaud, D., Moreno, C.: The theoretical, practical, and technological foundations of the 15-minute city model: proximity and its environmental, social and economic benefits for sustainability. Energies (Basel) **15** (2022). https://doi.org/10.3390/en15166042
5. Moreno, C., Allam, Z., Chabaud, D., Gall, C., Pratlong, F.: Introducing the "15-minute city": sustainability, resilience and place identity in future post-pandemic cities. Smart Cities **4**, 93–111 (2021). https://doi.org/10.3390/smartcities4010006
6. Mahmoudpour, A., Shirazi, M.R.: Justice at the core: Rethinking the 15-min city for just urban futures. Appl. Geography **183** (2025). https://doi.org/10.1016/j.apgeog.2025.103745
7. Fainstein, S.S.: The just city. Int. J. Urban Sci. **18**, 1–18 (2014). https://doi.org/10.1080/12265934.2013.834643
8. Chan, J.K.H.: Urban Ethics in the Anthropocene: the Moral Dimensions of Six Emerging Conditions in Contemporary Urbanism. Springer, Switzerland (2018)
9. Marcuse, P.: Spatial justice: derivative but causal of social injustice. Spatial Just. **1**, 1–6 (2009)
10. Lefebvre, H.: Le droit à la ville. L'Homme et la société. **6**, 29–35 (1967)
11. Harvey, D.: Social Justice and the City. University of Georgia Press (1973)
12. Nordquist, M.: Seeking spatial justice by Edward W. Soja. Contemporary Polit. Theo. **12**, e16–e18 (2013). https://doi.org/10.1057/cpt.2011.40
13. Iveson, K.: Social or spatial justice? marcuse and Soja on the right to the city. City **15**, 250–259 (2011). https://doi.org/10.1080/13604813.2011.568723
14. Soja, E.W.: Seeking Spatial Justice. U of Minnesota Press (2013)
15. Bednarowska-Michaiel, Z.: Ethnic inequalities in cycling to work in London: mobility injustice and regional approach. Reg. Stud. Reg. Sci. **10**, 475–488 (2023). https://doi.org/10.1080/21681376.2023.2186802

16. Annunziata, A., Murgante, B.: Unveiling the 15 minute anti-city: exploring spatial injustice in the context of urban e-planning. In: Nunes Silva, C. (ed.) Recent Advances and Prospects in Urban E-Planning, pp. 323–348. IGI Global Scientific Publishing, New York (2025). https://doi.org/10.4018/979-8-3693-5956-3.ch013
17. Murgante, B., Valluzzi, R., Annunziata, A.: Developing a 15-minute city: evaluating urban quality using configurational analysis. The case study of Terni and Matera, Italy. Appl. Geography. **162**, 103171 (2024). https://doi.org/10.1016/j.apgeog.2023.103171
18. Murgante, B., Patimisco, L., Annunziata, A.: Developing a 15-minute city: a comparative study of four Italian Cities-Cagliari, Perugia, Pisa, and Trieste. Cities **146**, 104765 (2024). https://doi.org/10.1016/j.cities.2023.104765
19. Zhang, J.: Inequalities in the quality and proximity of green space exposure are more pronounced than in quantity aspect: evidence from a rapidly urbanizing Chinese city. Urban For. Urban Green. **79**, 127811 (2023). https://doi.org/10.1016/j.ufug.2022.127811
20. Logan, T.M., Hobbs, M.H., Conrow, L.C., Reid, N.L., Young, R.A., Anderson, M.J.: The x-minute city: measuring the 10, 15, 20-minute city and an evaluation of its use for sustainable urban design. Cities **131**, 103924 (2022). https://doi.org/10.1016/j.cities.2022.103924
21. Khavarian-Garmsir, A.R., Sharifi, A., Hajian Hossein Abadi, M., Moradi, Z.: From garden city to 15-minute city: a historical perspective and critical assessment (2023). https://doi.org/10.3390/land12020512
22. C40 Cities Climate Leadership Group: C40 Mayors' Agenda for a Green and Just Recovery (2020). https://www.c40knowledgehub.org/s/article/C40-Mayors-Agenda-for-a-Green-and-Just-Recovery?language=en_US
23. Miller, D.: Mayors engaging with the G20: make COVID-19 recovery green and just to deliver a successful COP26 (2021)
24. Abdelfattah, L., Deponte, D., Fossa, G.: The 15-minute city: interpreting the model to bring out urban resiliencies. New scenarios for safe mobility in urban areas. In: Proceedings of the XXV International Conference Living and Walking in Cities (LWC 2021), 9–10 September 2021, Brescia, Italy, vol. 60, pp. 330–337 (2022). https://doi.org/10.1016/j.trpro.2021.12.043
25. Allam, Z., Bibri, S.E., Chabaud, D., Moreno, C.: The '15-Minute City' concept can shape a net-zero urban future. Humanit. Soc. Sci. Commun. **9**, 126 (2022). https://doi.org/10.1057/s41599-022-01145-0
26. Guzman, L.A., Oviedo, D., Cantillo-Garcia, V.A.: Is proximity enough? a critical analysis of a 15-minute city considering individual perceptions. Cities **148**, 104882 (2024). https://doi.org/10.1016/j.cities.2024.104882
27. Murgante, B., Annunziata, A.: Application of the 15-minute city criteria to a metropolitan area: a case study of the metropolitan city of Cagliari, Italy. Int. J. E-Planning Res. **14** (2025). https://doi.org/10.4018/IJEPR.371757
28. Perry, C.: The neighborhood unit. In: The City Reader, pp. 607–619. Routledge (2015)
29. Padeiro, M., Louro, A., da Costa, N.M.: Transit-oriented development and gentrification: a systematic review. Transp. Rev. **39**, 733–754 (2019). https://doi.org/10.1080/01441647.2019.1649316
30. Calthorpe, P.: The Next American Metropolis. Ecology, Community and the American Dream. Princeton Architectural Press, New York (1993)
31. Katz, P.: The New Urbanism: Toward an Architecture of Community. McGraw-Hill, Incorporated, New York, United States (1994)
32. Duany, A., Plater-Zyberk, E.: The second coming of the American small town. Wilson Q. **16**, 3–51 (1992)
33. Yamu, C., Frankhauser, P.: Spatial accessibility to amenities, natural areas and urban green spaces: using a multiscale, multifractal simulation model for managing urban sprawl. Environ. Plann. B. Plann. Des. **42**, 1054–1078 (2015). https://doi.org/10.1068/b130171p

34. Jacobs, J.: The Death and Life of Great American Cities. Vintage, New York (2016)
35. Alexander, C.: A Pattern Language: Towns, Buildings, Construction. Oxford University Press, Oxford (1977)
36. Gehl, J.: Life between Buildings: Using Public Space. Island press, Washington DC (2011)
37. Mahmoudpour, A., Shirazi, M.R.: Critical debates on the 15-minute city: a systematic content analysis of the literature. Cities **169** (2026). https://doi.org/10.1016/j.cities.2025.106576
38. Mouratidis, K.: Time to challenge the 15-minute city: Seven pitfalls for sustainability, equity, livability, and spatial analysis. Cities **153**, 105274 (2024). https://doi.org/10.1016/j.cities.2024.105274
39. Pirmoradi, P., Garshasbi, D., Solgi, N.: A critical review: the 15-minute city. Plan. Pract. Res. (2025). https://doi.org/10.1080/02697459.2025.2507334
40. Casarin, G., MacLeavy, J., Manley, D.: Rethinking urban utopianism: the fallacy of social mix in the 15-minute city. Urban Stud. **60**, 3167–3186 (2023). https://doi.org/10.1177/00420980231169174
41. ISTAT (Italian National Institute for Statistics): Popolazione e Famiglie (2023). https://esploradati.istat.it/databrowser/#/it/dw/categories
42. Mara, F., Deri, F., Anselmi, C.: Mismapped cities: the risks of using big-data in urban planning. In: De Lotto, R., Venco, E., Pietra, C. (eds.) Innovation in Urban and Regional Planning. INPUT 2025. Lecture Notes in Civil Engineering, vol. 826. Springer, Cham (2026). https://doi.org/10.1007/978-3-032-18822-9_38
43. Autorità per le Garanzie nelle Comunicazioni (AGCOM): Broadband Map Open Data. https://geo3.agcom.it/opendata/default.aspx. Accessed 23 Mar 2026
44. Vale, D., Lopes, A.S.: Accessibility inequality across Europe: a comparison of 15-minute pedestrian accessibility in cities with 100,000 or more inhabitants. NPJ Urban Sustain. **3**, 55 (2023). https://doi.org/10.1038/s42949-023-00133-w
45. Lord, A., Cheang, C.W., Dunning, R.: Understanding the geography of affordable housing provided through land value capture: evidence from England. Urban Stud. **59**, 1219–1237 (2022). https://doi.org/10.1177/00420980211998893
46. Pozoukidou, G., Chatziyiannaki, Z.: 15-minute city: decomposing the new urban planning Eutopia. Sustainability **13** (2021). https://doi.org/10.3390/su13020928
47. Bonfantini, G.B., Galimberti, B., Ventura, E.: There are no 'solutions' in urban planning: against the idea of a ready-made urbanism and the 15-minute city's uncritical branding. Land (Basel) **14** (2025). https://doi.org/10.3390/land14102090
48. Calafiore, A., Dunning, R., Nurse, A., Singleton, A.: The 20-minute city: an equity analysis of Liverpool City Region. Transp. Res. D Transp. Environ. **102**, 103111 (2022). https://doi.org/10.1016/j.trd.2021.103111

The Limits of the 15-Minute City: a Systematic Qualitative Literature Review

Beniamino Murgante and Alfonso Annunziata(✉)

Department of Engineering, University of Basilicata, 85100 Potenza, Italy
alfonso.annunziata@unibas.it

Abstract. The 15MC has gained prominence as a planning paradigm aimed at addressing spatial inequality, car dependency, and the environmental and social externalities of urban dispersion through proximity-based urban organization. Yet, despite its increasing influence in academic and policy debates, the concept remains marked by theoretical ambiguity, methodological inconsistency, and implementation challenges. This article presents a structured qualitative literature review of the limitations associated with the 15MC, organized via a conceptualization–operationalization–implementation–impact framework. The study combines systematic review procedures with AI-assisted semantic relevance ranking and inductive content analysis. A Scopus-derived corpus of 242 publications was first prioritized through semantic ranking, after which the 25 most relevant articles were qualitatively examined. The findings develop a systematic taxonomy of the concept's limitations, underlining conceptual tensions, measurement issues, governance constraints, and potential unintended effects. The article concludes by identifying key directions for reinforcing the model's theoretical, methodological, and policy foundations.

Keywords: 15MC · Limit · Critique · Policy

1 Introduction

In recent years the 15-min city (15MC) concept has emerged as a significant paradigm for addressing emerging urban issues related to spatial inequalities, environmental and climate crises, urban dispersion, and car-dependency [1–4]. Grounded in the principles of chrono-urbanism, the 15MC proposes an urban configuration in which six core functions - housing, employment, education, healthcare, commerce, and leisure - are located within a distance from each residence equivalent to a travel time of 15 min by foot or by bicycle [4–7]. The aim is to reduce travel time, promote active mobility, and foster proximity-based living. Chrono-urbanism emphasizes the spatio-temporal organization of urban systems, advocating for decentralization and re-distribution of functions to optimize individuals' time use. By minimizing travel time related to basic activities, this paradigm reallocates time toward social interaction and recreational activities. Moreover, reduced travel demand contributes to decreased car dependency, improved access to services of general interest (SGI), and increased reliance on active modes of transport.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 39–59, 2027.
https://doi.org/10.1007/978-3-032-30530-5_3

Adopting a time-centric planning perspective yields multiple co-benefits. These include more equitable access to urban amenities, lower energy consumption, reduced noise and air pollution, social cohesion, and overall improvements in urban quality of life [4, 5]. This concept has emerged as a significant interpretative lens for evaluating cities' performance in terms of access to essential functions [1, 8–10], and as a model informing urban policies and planning initiatives [11, 12]. At the same time, the model has been subject to substantial criticism regarding its conceptual foundations, issues of replicability and transferability, structural impediments to the development of robust and standardized compliance assessment protocols, feasibility and governance constraints, and the potential for unintended adverse effects. This article presents a structured qualitative literature review of the limitations, critiques, and objections concerning the 15MC concept, organized via a conceptualization–operationalization–implementation–impact lens. The study combines systematic literature review and inductive content analysis, by introducing a procedure that integrates computational and qualitative methods. First, it uses semantic relevance ranking of a Scopus-derived corpus (n = 242) to prioritize the literature. Then, an inductive content analysis is conducted on the 25 most relevant articles. By combining AI-assisted retrieval techniques and comprehensive qualitative synthesis, the study advances a systematic taxonomy of the limitations of the 15MC concept. This article contributes to the debate on the 15MC in three ways. First, it develops a systematic taxonomy of the concept's limitations across four analytical dimensions - conceptualization, operationalization, implementation, and impact - clarifying the structural, methodological, and governance-related barriers that may prevent urban policies based on the model from delivering expected outcomes in terms of equity, social inclusion, sustainability, and health. Second, it advances methodological practice by combining AI-assisted semantic relevance ranking with inductive content analysis, enhancing transparency and replicability while maintaining interpretive rigor. Third, it identifies priority inquiry directions to improve the model's conceptual foundations, measurement procedures, and policy feasibility, thus contributing to bridging the gap between normative urban visions and operational planning tools. The article is structured into six sections. After the introduction, Sect. 2 outlines the conceptual foundations, and Sect. 3 describes the proposed method. Section 4 presents the results, which are discussed in Sect. 5. Finally, the Conclusions summarize the main findings and outline directions for future research.

2 Theoretical Framework and Key Concepts

The 15-min city is defined as an urban model based on chrono-urbanism aimed at ensuring that all residents are able to access all basic essential functions - including Living, Employment, Commerce, Healthcare, Education, and Entertainment - within 15 min on foot or by bicycle. The model is based on the principle that the quality of urban life is inversely proportional to time spent in automobile-based commuting and aims to counteract the socio-spatial fragmentation produced by modernist, car-dependent planning paradigms. More precisely, the 15MC aims to address interrelated issues, including environmental degradation, climate alteration, spatial and social inequalities, declining public health, and inadequate urban resilience exposed by the COVID-19 pandemic. In

this sense, the concept advances a proximity-based planning concept intended to improve urban resilience, sustainability, and social cohesion while reducing car dependency and improving equitable access to essential services.

The 15-min city concept entails a reorganization of cities compliant with the components of density, proximity, diversity, and digitalization. Density is framed as a functional and relational dimension, and refers to the optimal number of people per unit area that can be sustained in terms of service provision and resource consumption [4, 5]. A complementary definition considers density as the optimal number of residents per unit area required to sustain service diversity and plurality [6, 7]. Proximity is conceptualized as a multidimensional principle integrating spatial and temporal access to essential urban functions. Proximity refers to the close spatial distribution, decentralization, and redistribution of services at the district scale and promotes the multifunctional use of urban infrastructures [4, 6, 7, 13]. It supports active mobility, reduces automobile dependence, enhances inclusiveness and social interaction, and contributes to urban sustainability and quality of life [14–16]. Diversity refers to the coexistence and integration of varied urban functions and socio-cultural identities at the local scale [4, 5, 7, 17]. It encompasses the presence of mixed-use environments and the co-presence of heterogeneous populations, practices, and cultural expressions. Lastly, digitalization, or digital ubiquity, refers to the use of Information and Communication Technologies (ICT), Internet of Things (IoT) systems, distributed sensors, and digital twin platforms to optimize service delivery, foster inclusivity, enhance participation, and generate data streams that can be analyzed to inform decision-making, measure performance indicators, and monitor policy implementation and outcomes [4, 6, 18, 19].

The 15MC concept, particularly in the aftermath of the COVID-19 pandemic, has emerged as a robust interpretative lens for assessing urban performance in terms of equitable access to essential services and functions. It has also been advanced as a policy model for addressing environmental degradation, climate alteration, and socio-spatial inequalities. Consequently, a growing body of literature proposes quantitative methods to evaluate the extent to which urban areas align with the foundational principles of the 15MC. Empirical applications include assessments of equitable access to urban functions in the Liverpool City Region [20], in Barcelona [21], Oslo and Lisbon [22] in the Italian cities of Terni, Matera [6], Cagliari, Perugia, Pisa, Trieste [7]. Comparative studies measure cumulative opportunities and variety of destinations across European Cities with more than 100,000 inhabitants [23], compliance of distinct city typologies with 15MC criteria across 24 global cities [24], and measure levels of proximity across the 500 largest cities in the United States and 43 urban areas in New Zealand [25]. Moreover, the C40 initiative - comprising 100 major cities across multiple countries - has identified the 15MC concept as a relevant paradigm for sustainable and equitable post-pandemic recovery [26]. Numerous global cities - including Paris, New York, Washington, Dublin, London, Seattle, Portland, Toronto, Melbourne, Bogotá, Milan, Lisbon, and Rome - have incorporated the 15MC concept into their urban policy documents as a model to improve proximity, ensure equitable access to services of general interest, and foster social cohesion [6, 7]. Despite its prominence in academic and policy discourse, substantial criticism has emerged regarding the conceptualization, operationalization, practical feasibility, and unintended consequences of the 15MC model. In response to

these debates, the subsequent sections present a structured qualitative literature review that aims to systematize a taxonomy of the principal limitations associated with the 15MC concept.

3 Methodology

The analysis integrates a structured literature review with an inductive qualitative content analysis. The objective is to develop a systematic taxonomy of the limitations, objections, and critiques of the 15MC concept, organized along the dimensions of conceptualization, operationalization, implementation, and impact. The methodological procedure is structured into three stages: (i) corpus construction; (ii) semantic relevance ranking of the corpus; and (iii) inductive content analysis.

3.1 Corpus Construction

The corpus was constructed through a replicable search strategy applied to the Scopus database. Scopus was selected based on two criteria: (i) its extensive coverage of peer-reviewed scientific publications relevant to the research topic; and (ii) the completeness and standardization of its metadata, which facilitate subsequent bibliometric analysis.

Relevant articles focusing on the limitations of the 15MC concept were identified via the development and iterative refinement of search queries incorporating the terms "15-min city," "x-minute city," "limitation*," and "critic*." The preliminary query was subsequently refined to improve precision and recall. This refinement involved restricting the search to the TITLE-ABS-KEY field and incorporating frequently used terminological variants of the concept. The final search string was structured as follows:

TITLE-ABS-KEY (("15MC" OR "15 min city" OR "15min city" OR "20-min city" OR "20 min city" OR "20min city" OR "20-min neighbourhood" OR "20-min neighborhood" OR "15-min neighbourhood" OR "15-min neighborhood" OR ("N-minute" W/1 (city OR cities OR neighbourhood OR neighborhood)) OR ((minute* W/2 (city OR cities OR neighbourhood OR neighborhood)) AND (15 OR 20))) AND (limitation* OR critic* OR criticism OR drawback* OR challenge* OR concern*)). The structured search resulted in the construction of a corpus of 242 items.

3.2 Semantic Relevance Ranking

To prioritize the literature in relation to the research question, a semantic relevance ranking procedure was implemented. Rather than relying exclusively on Boolean keyword matching, the study adopted a transformer-based semantic similarity approach.

For each record, a textual representation was constructed by concatenating the title, abstract, author keywords, and index keywords.

These document representations were encoded into dense vector embeddings using a pre-trained Sentence-Transformers model (all-mpnet-base-v2). The embeddings were L2-normalized, and document relevance was calculated as the cosine similarity between each document vector and a natural-language query explicitly formulated to capture discussions of limitations, criticisms, and implementation challenges associated with the 15MC concept. The natural-language query was formulated as follows:

Table 1. List of scientific articles selected for the inductive content analysis

AU	TI	TC	S_SC
Mahmoudpour A; Shirazi Mr [27]	Critical Debates On The 15MC: A Systematic Content Analysis Of The Literature	2	0.827
Megahed G; Elshater A; Abdelmoneim Elrefaie M; Afifi S [11]	Exploring The X-Minute City: A Chronological Analysis And Thematic Review Of Proximity-Based Urbanism	0	0.815
Mouratidis K [28]	Time To Challenge The 15MC: Seven Pitfalls For Sustainability, Equity, Livability, And Spatial Analysis	69	0.808
Khavarian-Garmsir Ar; Sharifi A; Sadeghi A [29]	The 15MC: Urban Planning And Design Efforts Toward Creating Sustainable Neighborhoods	251	0.795
Khavarian-Garmsir Ar; Sharifi A; Hajian Hossein Abadi M; Moradi Z [30]	From Garden City To 15MC: A Historical Perspective And Critical Assessment	83	0.779
Allam Z; Khavarian-Garmsir Ar; Lassaube U; Chabaud D; Moreno C [2]	Mapping The Implementation Practices Of The 15MC	33	0.775
Iqbal A; Nazir H; Qazi Aw [31]	Exploring The 15-Minutes City Concept: Global Challenges And Opportunities In Diverse Urban Contexts	4	0.765
Pozoukidou G; Angelidou M [19]	Urban Planning In The 15MC: Revisited Under Sustainable And Smart City Developments Until 2030	79	0.764
Teixeira Jf; Silva C; Seisenberger S; Büttner B; et al. [32]	Classifying 15-Minute Cities: A Review Of Worldwide Practices	58	0.764
Arias-Molinares D; Geurs K; Grigolon A; Ulak B; et al. [33]	Exploring The 15MC Concept For The Urban Outskirts: A Systematic Literature Review	0	0.760
Lu M; Diab E [34]	Developing A 15MC Policy? Understanding Differences Between Policies And Physical Barriers	11	0.760
Carvalho T; Farber S; Manaugh K; El-Geneidy A [35]	Assessing The Readiness For 15-Minute Cities: A Literature Review On Performance Metrics And Implementation Challenges Worldwide	10	0.757

(*continued*)

Table 1. (*continued*)

AU	TI	TC	S_SC
Chau H-W; Gilzean I; Jamei E; Palmer L; Preece T; Quirke M [1]	Comparative Analysis Of 20-Minute neighborhood Policies And Practices In Melbourne And Scotland	22	0.751
Wang H; Tsoi Kh; Loo Bpy [1]	An Assessment Framework For 15-Minute Cities: Progress Worldwide And The Impact Of Urban Form	5	0.750
Pirmoradi P; Garshasbi D; Solgi N [36]	A Critical Review: The 15MC	1	0.750
Casarin G; Macleavy J; David Manley D [37]	Rethinking Urban Utopianism: The Fallacy Of Social Mix In The 15MC	60	0.743
Thornton Le; Schroers R-D; Lamb Ke; Daniel M; et al. [38]	Operationalising The 20-Minute neighborhood	61	0.731
Szymańska D; Szafrańska E; Korolko M [39]	The 15MC: Assumptions, Opportunities And Limitations	6	0.729
Sharma A; Chandi A; Winters M [40]	An Investigation Of 15-Minute neighborhoods In Surrey, British Columbia: A Community-Informed Social Equity Analysis For A Fast-Growing, Diverse, Canadian City	2	0.728
Sepehri B; Sharifi A [41]	X-Minute Cities As A Growing Notion Of Sustainable Urbanism: A Literature Review	24	0.727
Bonfantini Gb; Galimberti B; Ventura E [42]	There Are No 'Solutions' In Urban Planning: Against The Idea Of A Ready-Made Urbanism And The 15MC's Uncritical Branding	1	0.726
Baldwin J [43]	Proximity Without Equity? Racialized Housing Organizations And The 15MC	0	0.726
Papadopoulos E; Sdoukopoulos A; Politis I [9]	Measuring Compliance With The 15MC Concept: State-Of-The-Art, Major Components And Further Requirements	88	0.723
Branco Da Silva J; Ricardo Da Costa A; Morais De Sá A [44]	The 15MC: Application To Two Parishes Of The City Of Lisbon	1	0.722
Pinto F; Akhavan M [12]	Reimagining Proximity: Operationalising The X-Minute City And Urban Regeneration In Amsterdam And Milan	0	0.721

> "What are the limitations, criticisms, and implementation challenges of the 15MC (or 20-minute neighborhood) concept, including issues related to definition, equity, exclusion, gentrification, governance, scalability, measurement, and unintended consequences?"

While this formulation may have oriented the retrieval process towards these themes, the subsequent inductive coding did not impose these categories a priori, enabling additional or differently framed critiques to emerge from the selected set. Articles were subsequently ranked in descending order of similarity and exported for screening and further qualitative analysis.

3.3 Inductive Content Analysis

In the final stage, the 25 most semantically similar articles were selected for inductive content analysis (see Table 1). This threshold was adopted as a purposive and exploratory cut-off, suitable for a preliminary application of the analytical procedure and for enabling detailed qualitative coding of the most relevant contributions. During corpus screening, substantial thematic saturation was observed, as additional high-ranking records largely reiterated critique domains already identified in the selected set. Subsequently, the limitations and critiques of the 15MC model were systematically extracted and synthesized. The ultimate objective was to develop a taxonomy structured around four analytical dimensions: conceptualization, operationalization, implementation, and impact. More specifically, the conceptualization dimension encompasses limitations related to the theoretical assumptions underpinning the definition of the 15MC concept. The operationalization dimension refers to limitations associated with the development and application of quantitative methods and indicators used to assess the degree of alignment between urban environments and the 15MC framework. The implementation dimension encompasses barriers related to the feasibility, democratic legitimacy, and transparency of urban policies intended to reorganize urban environments in accordance with the 15MC framework. The impact dimension, by contrast, concerns the potential unintended negative consequences arising from the implementation of the model. The inductive analysis is articulated in four steps. First, each article was analyzed independently to identify all explicit and implicit limitations, criticisms, risks, or constraints related to the 15MC concept. No cross-article merging was conducted at this stage, and conceptually distinct limitations were retained to preserve analytical granularity. For each limitation, a structured record was created including: a neutral label; a definition; the mechanism explaining its relevance; a provisional domain classification; spatial level (where identifiable); impacted groups (if specified); and a supporting evidence excerpt with page reference. No interpretive generalization beyond the source text was introduced. After every five articles, a controlled within-batch normalization was performed. Each batch resulted in a refined limitation list, an Article $\times$ Limitation comparison matrix, and a descriptive summary of emerging patterns. Following article-level extraction and batch-level harmonization, a global normalization across subsets of articles was conducted by merging synonymous or near-synonymous limitation labels, removing redundancies, and clarifying boundaries among overlapping concepts. Secondly, each normalized limitation was allocated a primary macro-category - Conceptual, Measurement/Operationalization,

Governance/Implementation, or Impact. Then, a hierarchical taxonomy (articulated on three levels) was derived via rule-based assignment to categories and sub-categories, complying with size constraints (≤25 categories; ≤ 45 sub-categories). An analytical definition was formulated for each taxonomy node.

4 Results

The search strategy implemented in the Scopus database retrieved a corpus of 242 documents. Semantic relevance ranking, based on cosine similarity scores, confirmed the pertinence of the 100 most relevant items selected for the exploratory bibliometric analysis. These items directly address the research question concerning critiques and limitations of the 15MC concept in the scientific literature. The similarity scores ranged from 0.64 to 0.83, indicating a high degree of thematic alignment.

The inductive content analysis of the selected subset of 25 articles returns 255 observed limitations or critiques of the 15MC. Most critiques concern the governance/implementation domain (62 occurrences), operationalization and measurement issues (60 observed limitations), unintended adverse impacts on social justice and equity (44 instances), and inconsistencies in the conceptualization of the model (34 instances). The final taxonomy comprises 15 categories and 34 sub-categories (see Table 2).

The most frequently supported limitation category is Equity and Exclusion, addressed by 21 articles and associated with the Impact domain. This category encompasses four sub-categories related to unintended negative consequences of the 15MC concept: accessibility inequality and uneven service distribution; insufficient consideration of disability, ability, and diverse needs; exclusion, rights, and stigmatization; and segregation, stratification, and spatial exclusion mechanisms. Collectively, these sub-categories underline the inadequacy of proximity alone in addressing disparities in access to urban functions between central and semi-central areas [28, 37]. They also reflect the limited consideration of diverse mobility and accessibility needs determined by class, disability, and race [8, 29]; the locational advantages and disadvantages produced by district-level policies [19, 43]; and the risk that social-mix strategies may reduce the visibility of poverty rather than reinforce social capital [37]. Together, these dynamics may contribute to the reproduction of spatial inequalities [37, 42, 45].

The Category "Metrics and Operationalization" refers to gaps in indicator frameworks and evaluation design, and to limitations in operational definitions and knowledge. These issues include the absence of a unified measurement framework [8, 9, 27, 42], the loss of relevant information due to aggregation methods [11], and persistent gaps in practical definitions, implementation strategies, experiential evidence, and context-sensitive needs in operationalizing the 15MC [32, 39]. Such limitations lead to constrained transparency and accountability in policy performance, reducing relevance for planning and equity assessment, limiting robust cross-case comparison, and ultimately restricting evidence-based implementation and comparative learning [2, 11, 27].

Concerning the governance and implementation dimension, limitations related to Capacity and Coordination emerge as particularly significant. These include delays and accountability gaps, inter-departmental and multi-level coordination limitations, and institutional capacity constraints. These issues underline the persistent gap between policy discourse and actual implementation [11]. In particular, prolonged administrative

Table 2. Taxonomy of limitations of the 15MC concept

Category	Sub-Category	Definition
Context & Scale	Context dependence and non-transferability	Implementation and meaning vary strongly by local conditions, making generalization difficult[27, 29]
Context & Scale	Scale ambiguity and level mismatch	Scale ambiguity refers to unclear and inconsistent specification of the spatial scale of the 15MC, leading to misalignment between district-level interventions and city- or metropolitan-level governance and planning [29]
Definition & Framing	Ambiguous concept definition and scope	Criticisms emerge, underlining concerns that foundational principles, particularly proximity, remain inadequately specified [42]
Definition & Framing	SGI scope and functional coverage	The 15MC may adopt an overly simplistic understanding of accessibility by focusing predominantly on spatial proximity [11]
Definition & Framing	Branding, rhetoric and utopian framing	The 15MC is regarded as overly utopian and abstract, inadequately reflecting the complexity of urban phenomena and functioning as a place-branding slogan lacking clear metrics, benchmarks, and operational criteria [27, 29]
Definition & Framing	Originality and novelty critique	The concept may overstate its originality, reframing and rebranding pre-existing ideas under a distinct label [28, 42]
Determinism & Reductionism	Overemphasis on decentralisation	Emphasis on decentralization risks reducing the 15MC's transformations to a simplistic polycentric model, disregarding urban center hierarchy and the benefits of centralizing specialized functions [27, 29]

(*continued*)

Table 2. (*continued*)

Category	Sub-Category	Definition
Determinism & Reductionism	Oversimplification of complex urban systems	Simplistic and rigid planning structures overlook the interdependencies among urban subsystems, thereby weakening cities' resilience to systemic challenges such as climate change, pandemics, and transportation crises [19, 29]
Determinism & Reductionism	Reductionism to proximity/time threshold	Proximity reductionism refers to reducing urban complexity to fixed time-distance metrics, neglecting contextual specificity, infrastructure quality, and sociodemographic diversity [1]
Determinism & Reductionism	Spatial/physical determinism	Refers to reducing 15MC outcomes to spatial proximity alone, while neglecting the social, economic, and governance conditions shaping urban life [2, 30]
Capacity & Coordination	Bureaucracy, delays and accountability gaps	The critique posits that 15MC realization is constrained by fragmented, inadequately accountable governance, where complex processes, limited statutory force, and digital divides obstruct implementation [12]
Capacity & Coordination	Coordination gaps (inter-departmental / multi-level)	The critique posits that 15MC implementation is hindered by siloed departments and multilevel regulatory complexity, which limit state capacity for integrated action [11, 12]
Capacity & Coordination	Institutional capacity and skills gaps	This critique posits that 15MC implementation is constrained by limited administrative capacity, staffing, and technical expertise in local institutions [11]

(*continued*)

Table 2. (*continued*)

Category	Sub-Category	Definition
Politics & Legitimacy	Participation, representation and legitimacy gaps	15MC planning often relies on uneven, episodic or tokenistic participation, constraining representation and democratic legitimacy [19, 42]
Politics & Legitimacy	Public resistance and contested politics	15MC measures, particularly car restrictions, can determine public opposition, conflict, and legitimacy disputes that obstruct implementation [11]
Regulation & Instruments	Regulatory rigidity	Car-oriented traffic rules and national and regional legal constraints institutionalise automobility, limiting local sustainable mobility transitions [11]
Resources & Feasibility	Funding and cost constraints	15MC implementation is limited by insufficient funding and rising development costs, delaying implementation and constraining maintenance [19]
Resources & Feasibility	Implementation feasibility in existing urban forms	Implementation feasibility of 15MC principles in existing urban forms is limited by rigid zoning, cost burdens, and land scarcity; in suburban areas, these constraints are compounded by low density and structural car dependency [27, 29, 30, 34]
Economic & Labor	Economic trade-offs and externalities	District-scale gains from 15MC policies may displace environmental and economic costs across scales, complicating net-impact assessment and causal attribution [27, 29, 30, 34]
Economic & Labor	Local jobs feasibility and agglomeration limits	Labour-market agglomeration and job concentration make universal local employment unrealistic, requiring robust regional transit [27]

(*continued*)

Table 2. (*continued*)

Category	Sub-Category	Definition
Equity & Exclusion	Accessibility inequality	15MC models overprivilege local proximity while neglecting public transport, thereby reproducing unequal access to regional opportunities [28]
Equity & Exclusion	Disability/ability and diverse needs	15MC models rely on standard walking/cycling assumptions, disregarding disability, age-related mobility differences, and heterogeneous daily accessibility needs [27, 28]
Equity & Exclusion	Limited impact on segregation/injustice	Proximity-based planning cannot by itself redress structural inequality, exclusion, or segregation, and may reproduce injustice via selective inclusion, digital divides, and paternalistic social-mix logics [37]
Equity & Exclusion	Spatial exclusion mechanisms	District-based service allocation can generate uneven locational advantage, thus reproducing intra-urban spatial disparities and territorial segregation [19]
Housing & Affordability	Gentrification and displacement	Proximity-led improvements can capitalise into land and housing values, increasing exclusionary pressures and displacing lower-income residents from areas with greater service provision [27, 31]
Mobility Outcomes	Limitations to modal transition	Proximity alone does not modify travel practices; without safe, continuous pedestrian and cycling infrastructure, mode transition and access to urban functions remain limited [9, 27, 44]

(*continued*)

Table 2. (*continued*)

Category	Sub-Category	Definition
Data & Evidence	Data availability, coverage, and quality gaps	15MC evaluation is constrained by incomplete, non-longitudinal, and low-granularity data, which obscure quality-of-life, amenity quality, and equity impacts [2]
Methodological & Scale Sensitivity	Composite index and weighting issues	15MC indices can be distorted by arbitrary weighting and equal treatment of opportunities, neglecting distance sensitivity and limiting measurement validity [40]
Methodological & Scale Sensitivity	Scale/aggregation sensitivity and counting problems	Operational specifications related to the definition of spatial units, aggregation procedures, and threshold specifications can generate MAUP-related distortions, edge effects, and unstable accessibility assessments [8]
Metrics & Operationalization	Indicator scope omissions	15MC metrics omit substantive dimensions - including amenity quality, public transport, digital access, and equity - yielding incomplete and policy-misleading assessments [9, 34]
Metrics & Operationalization	Benchmarks/targets unclear	15mC assessment relies on non-standardised operational definitions, measurement frameworks, accessibility parameters, and temporal assumptions, hence reducing cross-context and cross-group comparability and constraining evaluative robustness and accountability [27]

(*continued*)

Table 2. (*continued*)

Category	Sub-Category	Definition
Metrics & Operationalization	Operational definitions and knowledge gaps	A Limited agreed evidence base on definitions, implementation experiences, and needs and issues in operationalising 15mC across contexts constrains evidence-based action and comparative learning [32]
Metrics & Operationalization	Improper aggregation of facilities	Improper facility aggregation conflates distinct essential services into composite accessibility scores, obscuring service-specific deficits and biasing proximity assessment [28]
Thresholds & Standards	15-min threshold validity and time budgets	Fixed x-minute cut-offs provide insufficient behavioural and contextual justification, rendering 15mC assessments sensitive to the selected time limit, of limited comparability, and normatively fragile [27]

procedures and bureaucratic delays may undermine timely delivery [12], while siloed organizational structures and inadequate coordination across agencies and levels of government can obstruct coherent and integrated action, reducing the effectiveness of urban strategies [12, 39]. Moreover, local institutions may not possess sufficient personnel, specialized expertise, or administrative capability to adequately design, manage, and execute interventions consistent with the 15MC concept [11, 19, 31, 45]. These limitations undermine the clarity, feasibility, and credibility of implementation processes. They may reduce progress, generate public dissatisfaction, decrease trust in policy commitments, and risk compromising the equitable distribution of the benefits generated by the implementation of the 15MC model.

Determinism and Reductionism emerge as significant limitations in the conceptualization of the 15MC. Physical determinism posits that spatial form and built-environment properties exert direct causal influence on individual practices and socio-economic outcomes [2, 27, 29, 30]. This perspective implies that complex social and economic problems can be solved via urban configuration and proximity-based interventions alone [27, 29]. An excessive focus on decentralization can lead to the risk of reducing the complex social, environmental, and economic transformations envisioned by the 15MC model into a simplistic image of a polycentric city [27, 29]. Moreover, extreme decentralization disregards the hierarchical structure of urban centers and services [28], and the advantages of centralizing specialized services and functions [39]. Lastly, the decentralization of labor

may conflict with prevailing economic dynamics that favor hyper-specialisation [27, 30]. A related aspect concerns the risk of oversimplification of complex structures: simplistic, rigid structures disregard the interdependencies among urban sub-systems, thus reducing cities' resilience to climate alteration, pandemics, and transportation crises [29]. Moreover, proximity reductionism refers to the tendency to oversimplify the complexities of urban and social life, disregarding the interactions among complex sub-systems, hence compressing urban planning to the sole, variously interpretable dimension of spatial proximity [27, 36]. Moreover, Uniform targets, in terms of time and distance, disregard context specificity and needs subjectivity, reducing feasibility and meaningfulness of evaluation [1, 44]. Finally, critics argue that proximity alone is insufficient to guarantee equity and social cohesion [37]. Further critiques address an area at the intersection of conceptualization and governance implementation dimensions: definition and framing issues include limitations related to Ambiguous concept definition and scope, Scope of amenities and functions, Branding, rhetoric and utopian framing, and Originality and novelty critiques. The latter underlines the overstatement of the concept's originality, positing that the 15MC concept may represent a recirculation and reframing of preexisting ideas under a distinct brand name [28]. Regarding the scope of amenities and functions, critics argue that an exclusive emphasis on distance may neglect other critical dimensions of access, including availability, acceptability, affordability, adequacy, and awareness of services [1, 40]. This reductionist perspective can undermine the validity of 15MC classifications and result in misguided or incomplete policy interventions [1].

Ambiguity in concept definition and scope, refers to concerns that foundational principles, particularly proximity, remain inadequately specified [27, 28, 42]. Varying interpretations relating to targets, destination types, and thresholds - as well as the tendency to employ the concept as a broad guideline [33] rather than a framework integrated with specific measures, regulations, and implementation tools [34, 42] - contribute to conceptual uncertainty. Such ambiguities may result in inconsistent policy strategies and uneven implementation, ultimately limiting accountability, standardization, and comparability across cases [34]. Lastly, the 15MC concept is regarded as overly utopian and abstract, inadequately reflecting the complexity of urban phenomena and functioning as a place-branding slogan, unsubstantiated by clear metrics, benchmarks, and operational criteria [27, 33, 42].

5 Discussion

The results underline several critical issues that limit both the implementation of the 15MC concept and its transformative potential across distinct dimensions. Overall, the findings of the literature review provide relevant indications for policy design and future research. Concerning conceptualization, a major critical issue regards the reduction of the 15MC to a branding slogan, unsupported by clear metrics, operational criteria, and an organic urban strategy articulated via interdependent and trans-scalar actions [27, 33, 42]. In such cases, implementation may result in isolated and episodic interventions merely framed under the umbrella of the 15MC. This underscores the need for urban policies to recognize and incorporate the set of interdependent actions and processes of structural reorganization required to create the functional and spatial conditions necessary to sustain a minimum demand base and enable the relocation or redistribution of

urban functions [6, 7, 24, 46, 47]. These actions include the densification of semi-central areas, the intensification of the system of urban spaces and its reconfiguration into a more continuous grid through the regeneration of marginal spaces, and the improvement of enclosure conditions and urban porosity arising from the interaction between public and private spaces [7, 46]. Such interventions increase permeability, spatial integration, and the number of potential users, thus fostering the co-presence of pedestrians in public space. In line with the principle of movement economy, pedestrian co-presence constitutes a fundamental precondition for the emergence of centres and for the localization of services in the most vital urban areas [48]. A second, related concern lies in the predominant focus on spatial proximity, often to the neglect of the complementary conditions that inform access to services of general interest and the complexities of urban social and economic problems [27, 36]. This limitation points to the need for metrics measuring service quality, availability, and affordability. More broadly, it calls for analytical procedures able to incorporate multiple dimensions of justice, particularly recognitional and procedural justice, digital divide, and socio-economic conditions [13].

Pursuing unrealistic decentralization targets and outlining a simplified polycentric urban configuration result in inadequate consideration of the diversification of centres, the articulation of the urban structure across local and urban-scale centralities, the specific functional role of each centre, and the concentration of specialized functions in urban-scale centres [27]. As a result, it becomes necessary to define the amenities to be included, focusing decentralization on services of general interest that are accessed on a daily basis and meet everyday needs at the district scale. Access to public transport also remains a central component of distributional equity, as it connects districts to broader urban structures organized around centralities of different levels and functions, and ensures access to specialized urban-scale services [37].

More broadly, social justice concerns reveal the limitations inherent in operationalizing the 15MC by means of uniform standards and fixed limit values. This indicates the need to define distance- and time-based limit values differentiated by service category and user group. It also requires the incorporation into measurement procedures of the perceived importance of services, as determined by the needs, preferences, and abilities of different users [1].

Limitations concerning the conceptualization and operationalization of the 15MC model imply the development of articulated and multidimensional measurement tools for the identification and prioritization of underserved areas presenting multiple dimensions of socio-spatial injustice, including distributional, recognitional, and procedural injustice [13, 49].

Concerns related to implementation feasibility within existing urban forms, particularly in marginal and peri-urban areas, imply two distinct sets of considerations. From a planning perspective, they underline the need for interventions aimed at reinforcing and creating local centralities via densification around existing public services and the regeneration of amenities as multifunctional spaces incorporating social, administrative, cultural, and civic functions [50]. These interventions can increase proximity to urban resources. At the same time, the reinforcement of mobility-as-a-service solutions and demand-responsive transport can improve access to urban functions from areas of dispersed urbanization.

From a governance perspective, these concerns point to the need for public–private cooperation and the relevance of a two-level governance process, whereby a coherent urban-scale vision is implemented through local interventions in specific district contexts. This strategy entails prioritizing deprived and vulnerable areas and implementing district-scale regeneration via participatory laboratories and context-specific governance structures; the objective is to contribute to broader urban restructuring and address issues of scale ambiguity, participation, and legitimacy. This issue also intersects with the need for mitigating governance barriers and coordination gaps. A fundamental concern relates to the risks of direct displacement, exclusionary displacement pressure, and gentrification [51]. Urban regeneration interventions may reproduce forms of marginalization, particularly in contexts already affected by socio-spatial inequality [27, 31]. Addressing these dynamics requires the integration of anti-displacement measures into regeneration policy, including linearly decreasing property taxes around urban regeneration interventions, property tax subsidies targeted at vulnerable residents, zoning buffer policies limiting transactions in areas contiguous to intervention sites, and social housing measures [27, 52]. The latter may take the form of direct public provision in priority areas, mandatory or negotiated on-site affordable housing requirements in private redevelopment, and in-lieu contributions, including financial payments or land transfers. In combination, these instruments can promote social mix in regeneration areas and expand access to affordable housing. At the same time, urban regeneration strategies require participatory governance processes and co-design activities capable of incorporating the needs, perceptions, and lived experiences of residents, particularly those of vulnerable and underrepresented groups [13, 27]. Within this perspective, co-design, participation, and anti-displacement measures become essential to address gaps in participation, representation, and legitimacy, to improve equity, and to reduce the potential displacement effects associated with urban regeneration.

6 Conclusions

This study develops a taxonomy of the emerging limitations of the 15MC concept, structured around four interdependent dimensions: conceptualization, operationalization, implementation, and impact or unintended consequences. The article aims to advance the debate on the adaptation of urban environments to contemporary urban issues, including environmental degradation, climate alteration, and persistent socio-spatial inequalities. More precisely, the study identifies fundamental directions for future inquiry and derives policy implications for reducing the risk that reductionist interpretations, fragmented strategies, simplistic forms of operationalization, and inconsistent implementation may not only fail to deliver the expected outcomes in terms of equity, social inclusion, and sustainability but also reproduce existing inequalities.

The article's contribution relies on three main aspects. First, it systematizes the limitations of the 15MC via a structured taxonomy that clarifies the conceptual, structural, and governance-related barriers.

Second, it introduces a review procedure that combines AI-assisted semantic relevance ranking with inductive content analysis, thus increasing the transparency and replicability while preserving interpretive rigor. Third, it contributes to bridging the gap

between normative urban visions and operational planning by identifying the conditions under which the 15MC can be translated into coherent policy and planning practice. The findings also open several directions for future research. These include the refinement of the proposed taxonomy through further empirical testing; the review of urban plans and policy documents incorporating the 15MC model, in order to assess the limitations that emerge in its translation into policy documents; and the identification of measures capable of addressing governance and implementation barriers, as well as mitigating unintended consequences such as gentrification, displacement, the reproduction of marginalization among vulnerable social groups, and persistent intra-urban disparities.

References

1. Wang, H., Tsoi, K.H., Loo, B.P.Y.: An assessment framework for 15-minute Cities: progress worldwide and the impact of urban form. Transp. Res. Part A Policy Pract. **199** (2025). https://doi.org/10.1016/j.tra.2025.104583
2. Allam, Z., Khavarian-Garmsir, A.R., Lassaube, U., Chabaud, D., Moreno, C.: Mapping the implementation practices of the 15MC. Smart Cities **7**, 2094–2109 (2024). https://doi.org/10.3390/smartcities7040083
3. Marquet, O., Mojica, L., Fernández-Núñez, M.B., Maciejewska, M.: Pathways to 15MC adoption: can our understanding of climate policies' acceptability explain the backlash towards x-minute city programs? Cities **148** (2024). https://doi.org/10.1016/j.cities.2024.104878
4. Moreno, C., Allam, Z., Chabaud, D., Gall, C., Pratlong, F.: Introducing the "15MC": sustainability, resilience and place identity in future post-pandemic cities. Smart Cities. **4**, 93–111 (2021). https://doi.org/10.3390/smartcities4010006
5. Allam, Z., Bibri, S.E., Chabaud, D., Moreno, C.: The '15MC' concept can shape a net-zero urban future. Humanit. Soc. Sci. Commun. **9**, 126 (2022). https://doi.org/10.1057/s41599-022-01145-0
6. Murgante, B., Valluzzi, R., Annunziata, A.: Developing a 15MC: evaluating urban quality using configurational analysis. The case study of Terni and Matera, Italy. Appl. Geog. **162**, 103171 (2024). https://doi.org/10.1016/j.apgeog.2023.103171
7. Murgante, B., Patimisco, L., Annunziata, A.: Developing a 15MC: a comparative study of four Italian Cities-Cagliari, Perugia, Pisa, and Trieste. Cities **146**, 104765 (2024). https://doi.org/10.1016/j.cities.2023.104765
8. Carvalho, T., Farber, S., Manaugh, K., El-Geneidy, A.: Assessing the readiness for 15-minute cities: a literature review on performance metrics and implementation challenges worldwide. Transp. Rev. **45**, 801–827 (2025). https://doi.org/10.1080/01441647.2025.2513530
9. Papadopoulos, E., Sdoukopoulos, A., Politis, I.: Measuring compliance with the 15MC concept: State-of-the-art, major components and further requirements (2023). https://doi.org/10.1016/j.scs.2023.104875
10. Akrami, M., Sliwa, M.W., Rynning, M.K.: Walk further and access more! exploring the 15MC concept in Oslo, Norway. J. Urban Mob. **5** (2024). https://doi.org/10.1016/j.urbmob.2024.100077
11. Megahed, G., Elshater, A., Abdelmoneim Elrefaie, M., Afifi, S.: Exploring the X-minute city: a chronological analysis and thematic review of proximity-based urbanism. In: IOP Conference Series: Earth and Environmental Science. Institute of Physics (2025). https://doi.org/10.1088/1755-1315/1530/1/012020
12. Pinto, F., Akhavan, M.: Reimagining Proximity: Operationalising the X-Minute City and Urban Regeneration in Amsterdam and Milan. Land (Basel) **15**, 71 (2025). https://doi.org/10.3390/land15010071

13. Mahmoudpour, A., Shirazi, M.R.: Justice at the core: rethinking the 15-min city for just urban futures. Appl. Geog. **183** (2025). https://doi.org/10.1016/j.apgeog.2025.103745
14. Elldér, E., Haugen, K., Vilhelmson, B.: When local access matters: a detailed analysis of place, neighbourhood amenities and travel choice. Urban Stud. **59**, 120–139 (2022). https://doi.org/10.1177/0042098020951001
15. Khavarian-Garmsir, A.R., Sharifi, A., Sadeghi, A.: The 15MC: urban planning and design efforts toward creating sustainable neighborhoods. Cities **132**, 104101 (2023). https://doi.org/10.1016/j.cities.2022.104101
16. Mouratidis, K., Yiannakou, A.: COVID-19 and urban planning: built environment, health, and well-being in Greek cities before and during the pandemic. Cities **121**, 103491 (2022). https://doi.org/10.1016/j.cities.2021.103491
17. Biraghi, C.A., Ogut, O., Dong, T., Tadi, M.: CityTime: a novel model to redefine the 15MC globally through urban diversity and proximity. Urban Sci. **9** (2025). https://doi.org/10.3390/urbansci9020036
18. Allam, Z., Bibri, S.E., Jones, D.S., Chabaud, D., Moreno, C.: Unpacking the '15MC' via 6G, IoT, and digital twins: towards a new narrative for increasing urban efficiency, resilience, and sustainability. Sensors **22** (2022). https://doi.org/10.3390/s22041369
19. Pozoukidou, G., Angelidou, M.: Urban planning in the 15MC: revisited under sustainable and smart city developments until 2030. Smart Cities **5**, 1356–1375 (2022). https://doi.org/10.3390/smartcities5040069
20. Calafiore, A., Dunning, R., Nurse, A., Singleton, A.: The 20-minute city: an equity analysis of Liverpool City Region. Transp. Res. D Transp. Environ. **102**, 103111 (2022). https://doi.org/10.1016/j.trd.2021.103111
21. Ferrer-Ortiz, C., Marquet, O., Mojica, L., Vich, G.: Barcelona under the 15MC Lens: mapping the accessibility and proximity potential based on pedestrian travel times. Smart Cities **5**, 146–161 (2022). https://doi.org/10.3390/smartcities5010010
22. Di Marino, M., Tomaz, E., Henriques, C., Chavoshi, S.H.: The 15MC concept and new working spaces: a planning perspective from Oslo and Lisbon. Eur. Plan. Stud. **31**, 598–620 (2023). https://doi.org/10.1080/09654313.2022.2082837
23. Vale, D., Lopes, A.S.: Accessibility inequality across Europe: a comparison of 15-minute pedestrian accessibility in cities with 100,000 or more inhabitants. NPJ Urban Sustainab. **3**, 55 (2023). https://doi.org/10.1038/s42949-023-00133-w
24. Burke, J., Gras Alomà, R., Yu, F., Kruguer, J.: Geospatial analysis framework for evaluating urban design typologies in relation with the 15MC standards. J. Bus. Res. **151**, 651–667 (2022). https://doi.org/10.1016/j.jbusres.2022.06.024
25. Logan, T.M., Hobbs, M.H., Conrow, L.C., Reid, N.L., Young, R.A., Anderson, M.J.: The x-minute city: measuring the 10, 15, 20-minute city and an evaluation of its use for sustainable urban design. Cities **131**, 103924 (2022). https://doi.org/10.1016/j.cities.2022.103924
26. C40 Cities Climate Leadership Group: C40 Mayors' Agenda for a Green and Just Recovery (2020). https://www.c40knowledgehub.org/s/article/C40-Mayors-Agenda-for-a-Green-and-Just-Recovery?language=en_US
27. Mahmoudpour, A., Shirazi, M.R.: Critical debates on the 15MC: a systematic content analysis of the literature. Cities **169** (2026). https://doi.org/10.1016/j.cities.2025.106576
28. Mouratidis, K.: Time to challenge the 15MC: seven pitfalls for sustainability, equity, livability, and spatial analysis. Cities **153**, 105274 (2024). https://doi.org/10.1016/j.cities.2024.105274
29. Khavarian-Garmsir, A.R., Sharifi, A., Sadeghi, A.: The 15MC: urban planning and design efforts toward creating sustainable neighborhoods. Cities **132** (2023). https://doi.org/10.1016/j.cities.2022.104101
30. Khavarian-Garmsir, A.R., Sharifi, A., Hajian Hossein Abadi, M., Moradi, Z.: From Garden city to 15MC: a historical perspective and critical assessment (2023). https://doi.org/10.3390/land12020512

31. Iqbal, A., Nazir, H., Qazi, A.W.: Exploring the 15-minutes city concept: global challenges and opportunities in diverse urban contexts (2025). https://doi.org/10.3390/urbansci9070252
32. Teixeira, J.F., et al.: Classifying 15-minute cities: a review of worldwide practices. Transp. Res. Part A Policy Pract. **189** (2024). https://doi.org/10.1016/j.tra.2024.104234
33. Arias-Molinares, D., et al.: Exploring the 15MC concept for the urban outskirts: a systematic literature review (2025). https://doi.org/10.1186/s12544-025-00743-8
34. Lu, M., Diab, E.: Developing a 15MC policy? understanding differences between policies and physical barriers. Transp. Res. Part A Policy Pract. **191** (2025). https://doi.org/10.1016/j.tra.2024.104307
35. Chau, H.W., Gilzean, I., Jamei, E., Palmer, L., Preece, T., Quirke, M.: Comparative analysis of 20-Minute neighbourhood policies and practices in Melbourne and Scotland. Urban Plan. **7**, 13–24 (2022). https://doi.org/10.17645/up.v7i4.5668
36. Pirmoradi, P., Garshasbi, D., Solgi, N.: A critical review: the 15MC. Plan. Pract. Res. (2025). https://doi.org/10.1080/02697459.2025.2507334
37. Casarin, G., MacLeavy, J., Manley, D.: Rethinking urban utopianism: the fallacy of social mix in the 15MC. Urban Stud. **60**, 3167–3186 (2023). https://doi.org/10.1177/00420980231169174
38. Thornton, L.E., et al.: Operationalising the 20-minute neighbourhood. Int. J. Behav. Nut. Phys. Activity **19** (2022). https://doi.org/10.1186/s12966-021-01243-3
39. Szymańska, D., Szafrańska, E., Korolko, M.: The 15MC: assumptions, opportunities and limitations. Bull. Geography. Socio-economic Series. 137–151 (2024). https://doi.org/10.12775/bgss-2024-0038
40. Sharma, A., Chandi, A., Winters, M.: An investigation of 15-minute neighbourhoods in Surrey, British Columbia: a community-informed social equity analysis for a fast-growing, diverse, Canadian city. J. Transp. Health. **42** (2025). https://doi.org/10.1016/j.jth.2025.102021
41. Sepehri, B., Sharifi, A.: X-minute cities as a growing notion of sustainable urbanism: a literature review. Cities **161** (2025). https://doi.org/10.1016/j.cities.2025.105902
42. Bonfantini, G.B., Galimberti, B., Ventura, E.: There are no 'solutions' in urban planning: against the idea of a ready-made urbanism and the 15MC's uncritical branding. Land (Basel). **14** (2025). https://doi.org/10.3390/land14102090
43. Baldwin, J.: Proximity without equity? Racialized housing organizations and the 15MC, (2026). https://doi.org/10.1080/17549175.2026.2613396
44. Branco da Silva, J., Ricardo da Costa, A., Morais de Sá, A.: The 15MC: Application to two parishes of the city of Lisbon. Sustainability (Switzerland) **17** (2025). https://doi.org/10.3390/su17041461
45. Pozoukidou, G., Chatziyiannaki, Z.: 15MC: Decomposing the new urban planning Eutopia. Sustainability **13** (2021). https://doi.org/10.3390/su13020928
46. Murgante, B., Annunziata, A.: Application of the 15MC criteria to a metropolitan area: a case study of the metropolitan city of Cagliari, Italy. Int. J. E-Plann. Res. **14** (2025). https://doi.org/10.4018/IJEPR.371757
47. Moroni, S.: Urban density after Jane Jacobs: the crucial role of diversity and emergence. City, Territory Architec. **3** (2016). https://doi.org/10.1186/s40410-016-0041-1
48. Hillier, B.: Centrality as a process: accounting for attraction inequalities in deformed grids. Urban Des. Int. **4**, 107–127 (1999). https://doi.org/10.1080/135753199350036
49. Feitosa, F.O., Wolf, J.H., Lourenço Marques, J.: Operationalizing spatial justice in urban planning: bridging theory with practice. Urban Res. Pract. 1–17 (2024). https://doi.org/10.1080/17535069.2024.2341254
50. Romano, B., Zullo, F., Fiorini, L., Ciabò, S., Marucci, A.: Sprinkling: an approach to describe urbanization dynamics in Italy. Sustainability **9** (2017). https://doi.org/10.3390/su9010097

51. Friesenecker, M., Thaler, T., Clar, C.: Green gentrification and changing planning policies in Vienna? Urban Res. Pract. **17**, 393–415 (2024). https://doi.org/10.1080/17535069.2023.2228275
52. Mendonça, R., Roebeling, P., Fidélis, T., Saraiva, M.: Can policy instruments enhance the benefits of nature-based solutions and curb green gentrification? The case of Genova, Italy. Environ. Dev. **50** (2024). https://doi.org/10.1016/j.envdev.2024.100995

Main Streets or Latent Corridors? Exploring the Significance of Weak Ties in Suburban Contexts

Federico Mara(✉), Federica Deri, and Valerio Cutini

Department of Energy, Systems, Territory and Construction Engineering, University of Pisa, Largo Lucio Lazzarino, 56122 Pisa, Italy
federico.mara@ing.unipi.it

Abstract. The ongoing debate on proximity-based planning has produced a wide range of methods for assessing access to everyday services. Yet, most of these approaches remain centred on what can be reached within a given time threshold, while giving less attention to the spatial structure that makes some streets more strategically relevant than others. This paper argues that, particularly in suburban areas, streets should be evaluated for their structural strategic value. Special attention should be given to corridors that connect weakly linked sub-systems, as they could support future proximity-oriented interventions. Building on recent research about weak ties in spatial systems and the Kemeny-based centrality measure (KBC), this paper proposes a simple diagnostic approach to bridgeness, developed within the framework of the DUT-funded EMC2 project. The aim is not to compare alternative bridgeness metrics, but to examine whether high-KBC segments in a suburban area already coincide with local functions, or instead reveal structurally important yet underused corridors. The approach is tested in Turano, a suburban area of Massa, Italy, located within the wider Versilia conurbation. Using a road-centreline representation and a verified inventory of amenities, the study investigates the relationship between bridging structure and the distribution of Points of Interest through the concepts of alignment and mismatch. The results offer an exploratory but operational framework for distinguishing between corridors that already act as local functional anchors and latent spatial assets whose potential is not yet reflected in the current distribution of services. The paper concludes that bridgeness can provide a useful additional layer for proximity-oriented planning, particularly when combined with reliable local data and with broader assessments of walkability, public-space quality, and functional needs.

Keywords: weak ties · Kemeny-based centrality · suburban proximity · POIs · EMC2 · decision support system

1 Introduction

The current discussion about the X-minute city has significantly broadened the tools available to urban planners. Accessibility measures, isochrones, network-based evaluations, and proximity scores are now commonly used to evaluate whether daily services

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 60–74, 2027.
https://doi.org/10.1007/978-3-032-30530-5_4

can be reached with short travel times [1]. Recent research has shown that proximity is not an inherent characteristic of urban space. Instead, it is shaped by choices regarding scale, functions, datasets, thresholds, spatial representation, and assessment logic [2]. As a result, two studies can examine the same location and produce different "proximity landscapes" because of their modelling decisions. This issue is particularly relevant in suburban areas, where is crucial identifying where the spatial system might already have the potential for future change [3]. The EMC2 model, developed within the Driving Urban Transition framework, specifically addresses this by suggesting a pathway for transformation in metropolitan peripheries. It focuses on creating compact forms centred around pedestrian-friendly main streets, mixed uses, and stronger mobility connections [4]. From this perspective, suburban planning goes beyond merely checking for the presence or absence of services. It also needs to pinpoint where local structure, movement potential, and urban form may support or obstruct proximity-focused strategies. In this context, recent work in space syntax and network analysis has revived the concept of weak ties in spatial terms. In social network analysis, weak ties are important because they connect otherwise isolated communities and foster opportunities for sharing information [5]. In spatial systems, translating this idea must be done carefully, as spatial graphs do not support varying tie strengths like social networks do. Therefore, the focus shifts from "weakness" to the ability to bridge gaps: a spatial weak tie is defined not by the intensity of interaction but by its role in connecting loosely connected subsystems [3]. This thinking aligns with recent research on Kemeny-based centrality, which identifies a segment as crucial when its absence leads to a significant decline in navigability and inter-cluster connections [6]. This paper brings these two discussions together by focusing on the following question: do strategically important corridors in a suburban area align with the location of urban functions, or do they highlight underused spatial resources that could support future proximity-related actions within the EMC2 model? This question helps clarify if the current distribution of amenities matches the bridging logic of the suburban system, revealing where structural potential is already being used and where it is only partially tapped, thus examining if a bridge-sensitive configurational metric can help analyse local functional geography in suburban settings, moving beyond simple assessments of accessibility to identifying corridors that could be vital for future proximity-focused interventions. The study focuses on Turano, a locality in the municipality of Massa, in Tuscany, Italy. Turano does not represent an isolated neighbourhood but rather a portion of the Versilia conurbation system with the potential of X-minute city core within a system of interconnected local centres. In this context, a main street emerges from identifying a strong weak tie in the conurbated area, centred on Via Aurelia, supported by historical developments, relatively high local centrality, and considerable bridging capacity [7]. The area also shows a mix of challenges and potential: fragmented pedestrian and cycling infrastructure, lack of public spaces beyond parking lots, unsafe areas, and limited pedestrian activity despite encouraging layout features. Based on this, the paper creates a straightforward but useful diagnostic framework. High-KBC segments are identified as bridging corridors and then compared with local Points of Interest through a verified inventory, rather than relying solely on online data due to known issues with web-derived databases at the neighbourhood level. The core analytical distinction focuses on alignment, where strategic corridors correspond with

local functions, and mismatch, where high-KBC segments remain only partially active in functional terms. In the latter case, the corridor is viewed as latent: not because KBC is expected to predict land use, but because the relationship between spatial structure and functional presence seems incomplete. By connecting bridgeness to the actual layout of local functions, the paper provides a spatially grounded diagnostic approach to discussing suburban proximity than simple counts of services or threshold-based measures. The paper is structured as follows. Section 2 briefly discusses weak ties, bridgeness, and the importance of structure-function relationships in proximity planning. Section 3 presents the case study as a relevant example within the EMC2 framework. Section 4 outlines the data and analytical process. Section 5 discusses the findings through the concepts of alignment, mismatch, and latent corridors. Section 6 considers the contributions and limitations of the approach and concludes.

2 Weak Ties, Bridgeness, and Proximity-Oriented Planning

The translation of weak ties into spatial terms has been explored in recent work by Cutini [3]. The main takeaway is that we cannot interpret the original Granovetterian concept literally. Spatial adjacency is binary, not graded. Triadic closure does not function like it does in social networks. What we can transfer meaningfully is the logic of bridging. Some segments connect clusters, neighbourhood-like sub-systems, or weakly linked parts of the urban grid. This connection gives them an importance that goes beyond standard centrality measures. In earlier research, this spatial translation was developed by combining choice-like behaviour with bridge degree. This showed that some urban elements can act as structural-hole spanners, with these positions often aligning with literal bridges, city gates, bypasses, or other major connectors [3]. More recent work has formalized this logic through removal-based diagnostics. In this context, Kemeny-based centrality measures how much each segment contributes to the overall efficiency of diffusion processes by calculating the increase in navigational effort caused by the removal of a segment [6]. This is relevant for the current paper because proximity planning often focuses on accessibility outcomes while overlooking the internal structural role of streets. However, commercial and service activities do not distribute randomly, and instead depend heavily on location and respond to demand, accessibility, centrality, catchment logic, and the broader organization of movement in the street network [8–10]. Some functions thrive in dense, walkable areas, while others need larger catchments, corridor exposure, or economies of scale. The location of functions is partly shaped by structural opportunities in the urban grid. This does not mean that a configurational metric can predict land use in a certain way. Demand, ownership structure, regulation, business turnover, and local development histories all play key roles. However, spatial structure is not just a neutral backdrop [8]. A corridor linking otherwise distinct parts of a suburban area may be better equipped to support through-movement, visibility, and connections between neighbourhoods. Such corridors might already attract local businesses, or they may remain underused despite their structural importance. In either case, comparing these scenarios can be informative as it reveals whether structure and function reinforce each other and where spatial potential is only partially realized. From this viewpoint, bridgeness adds another layer to proximity-oriented planning. It does not

replace accessibility analysis, demand-side evidence, or the need to prioritize functions based on planning goals. Instead, it asks a different question: where does the suburban system already have structurally strategic corridors that could be especially important for future proximity-oriented actions? This insight not only helps identify where services are currently gathered but also indicates where the conditions for a stronger main-street logic might already be in place. In the EMC2 perspective, such corridors may represent the best opportunities for targeted reinforcement, selective densification, and improvements in public spaces or mobility that could support more effective suburban proximity conditions. They also guide public-private initiatives aimed at implementing and managing proximity-based neighbourhoods and local centres over the long term.

3 Case Study: Turano Within the Versilia Conurbation

The EMC2 project addresses suburban transition through an integrated reading of functions, urban forms, and implementation conditions. Within the DESTeC contribution, this perspective has been translated into a multiscalar workflow. Each local case is first positioned within the wider conurbation; it is then examined through diachronic, configurational, morpho-functional, and regulatory analyses; and it is finally interpreted in operational terms through four interrelated questions: where intervention potential is concentrated, which forms of transformation are spatially and functionally appropriate, how these transformations can be implemented within existing constraints, and who the relevant actors are in the process. Within this framework, Turano was selected as one of the local test areas for exploring the conditions under which a suburban environment may evolve towards a more viable proximity-based structure.

Turano is a suburban locality in the municipality of Massa, within the wider Versilia conurbation, and is considered here not as an isolated neighbourhood, but as a local node within a broader polycentric system. In the ongoing EMC2 work, the area emerged as particularly relevant because it lies along a corridor shaped by a strong weak-tie logic at the conurbation scale, centred on Via Aurelia and supported by historical patterns of linear densification, relatively high local centrality, and significant bridging capacity. This makes Turano especially suitable for the present paper: it is a case in which a potentially strategic suburban main-street structure is already detectable in configurational terms, while its urban performance remains only partially consolidated.

At the local scale, the area brings together many of the conditions that make suburban retrofitting both necessary and difficult: low-density settlement, dependence on motorised mobility, fragmented pedestrian and cycling continuity, limited public-space provision beyond parking areas, and insecure waiting conditions around bus stops. At the same time, it serves a socially meaningful population of about 2,359 inhabitants distributed across roughly 760 residential buildings, with around one quarter of residents aged over 65. This combination of structural relevance, functional incompleteness, and planning urgency makes Turano a particularly appropriate case through which to investigate whether bridge-sensitive corridors are already matched by local functions, or instead remain latent spatial assets (FIg. 1).

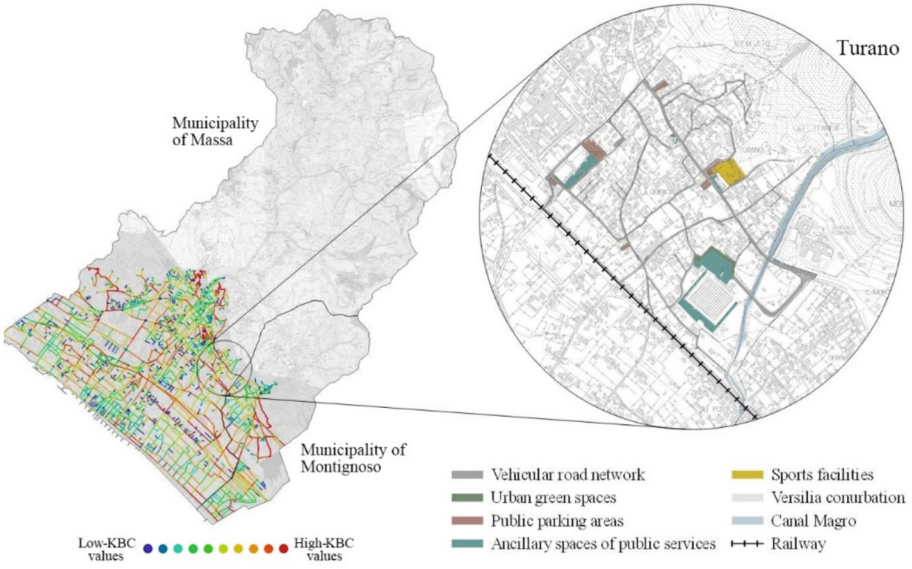


Fig. 1. Study area and local investigation boundary.

4 Data and Methods

4.1 Network Data and Spatial Representation

The analysis is based on a road-centre-line representation of the Turano street system, embedded within a substantially wider network extent. This choice is consistent with the broader configurational work developed within EMC2 and with recent studies on spatial weak ties, where segment-based network models are used to capture the relational structure of the urban grid and to identify corridors whose relevance derives not only from local prominence, but also from their role in connecting otherwise weakly coupled parts of the system.

Because KBC is highly sensitive to the structural embedding of each segment, the metric was not recalculated on a narrowly clipped local network. Instead, the pre-computed KBC values were derived from the wider Turano angular vehicular network and subsequently interpreted within the local study boundary. This approach reduces the risk of overestimating the importance of terminal or artificially clipped edges, and makes the local reading more robust with respect to boundary effects.

Within the Turano boundary, the network includes 482 segments. The analysis focuses on the upper tail of the local KBC distribution. The 90th percentile corresponds to a KBC value of 9,087.74, identifying 49 segments, namely the local top 10% higher-KBC set. These 49 segments do not form a single continuous line. Rather, they are organised into six disconnected clusters, which can be read more meaningfully as three main corridor systems: two interrupted stretches along Via delle Gorine, comprising 40 segments in total and about 929 m; three short clusters along Via Aurelia, comprising 7 segments and about 284 m; and a minor residual branch on Via del Papino, comprising 2 segments and about 22 m. This already indicates that bridgeness in Turano is not

diffusely distributed across the whole grid, but is concentrated in a limited number of structurally strategic alignments.

4.2 POI Data

The functional layer is based on a ground-truthed inventory of amenities produced through previous work on open-data reliability in suburban neighbourhoods [11–13]. This is a crucial methodological choice. Earlier analyses conducted on Turano and Bicchio showed that web-based datasets may underrepresent, overrepresent, or mislocate activities, with direct consequences for neighbourhood-scale proximity assessment. Using the field-verified layer in the present study therefore avoids reproducing these distortions, and ensures that the relation tested here is not between KBC and platform-dependent POIs, but between KBC and the most reliable available representation of actual local functions.

All functions are georeferenced and associated with a mesocategory, following the functional classification framework developed in recent work on proximity-oriented urban analysis. For the main purpose of this paper, however, POIs are initially treated as a single functional layer. This is consistent with the central aim of the study, namely to test the degree of alignment or mismatch between bridging structure and the current geography of local functions as a whole.

A more disaggregated reading by mesocategory remains possible and is considered here only as an interpretive extension, useful for discussing whether certain types of functions are more frequently associated with already activated corridors, while others are more common in latent ones. This allows the paper to retain a clear analytical focus, while still acknowledging that different amenity types may follow different spatial logics and planning priorities.

Within the study area, the inventory records 63 premises, of which 31 are vacant and 32 are occupied by commercial or service activities. Table 1 reports the aggregation of these activities into categories and mesocategories. The current distribution is already uneven: Via Aurelia alone contains 38 premises, equivalent to 60.3% of the total stock, including 18 occupied activities, or 56.3% of all active uses in the area. By contrast, Via delle Gorine contains only 3 premises, all of which are occupied. The rest of the local system is more fragmented, with smaller and more discontinuous concentrations along Via del Papino, Via San Remigio di Sotto, and a few minor streets.

Table 1. Local premises and activities by mesocategory.

Count	Meso-category	Categories
31	Vacant	vacant premises
5	Personal care	beauty salons, hairdressers, wellness centres
5	Private agencies	associations, event venues, waste/drainage services, law offices, psychotherapists
4	Animals and plants	florists, plant shops, agricultural supply stores

(*continued*)

Table 1. (*continued*)

Count	Meso-category	Categories
4	Food and beverages	bakeries and fresh pasta shops, greengrocers, beverage retailers
3	Bars and clubs	cafés, pastry shops, ice-cream
3	Vehicles	workshops, services, bicycle shops
2	Restaurant services	restaurants, pizzerias, trattorias
2	Furniture and home furnishings	home accessories and furniture stores
1	Clothing and accessories	clothing and sporting goods stores
1	Workshops and crafts	arts and crafts
1	Healthcare supplies and pharmacies	pharmacies and parapharmacies
1	Tobacco and tickets	tobacconists

4.3 KBC as Bridge-Sensitive Metric

Kemeny-based centrality is adopted here as the structural indicator because it is particularly well suited to the specific question addressed in the paper. Unlike conventional centrality measures that mainly capture gradients of accessibility or passage, KBC is explicitly sensitive to the contribution of each segment to the global coherence of the network. In operational terms, it identifies those streets whose removal would produce the greatest loss in stochastic navigability and inter-cluster coupling. This makes it especially appropriate when the aim is not simply to rank prominent streets, but to detect corridors that play a strategic bridging role within a suburban system. Since the paper is concerned with the relation between structural bridgeness and the geography of local functions, KBC provides a theoretically consistent and methodologically focused basis for the analysis.

Mathematically, let $G = (V, E)$ be the segment graph of the street network and let $P = D^{-1}A$ be the transition matrix of the natural random walk on that graph, where A is the adjacency matrix and D is the diagonal matrix of node strengths. The baseline network performance is summarised by the Kemeny constant K(G), which expresses the expected time needed for a random walker to reach a randomly chosen destination under the stationary distribution. For each segment $e = (i, j)$ in E, a perturbed graph G^(e) is constructed by removing the segment and compensating the operation with two self-loops at its endpoints: $G^(e) = (V, (E \setminus \{e\}) \cup \{(i, i), (j, j)\})$. The Kemeny-based centrality of segment e is then defined as $KBC(e) = K(G^(e)) - K(G)$. The larger the increase, the more that segment contributes to the overall efficiency and coupling of the network. In other words, segments with high KBC are those whose removal most strongly worsens the stochastic navigability of the system. In the present study, the field KBC_Simax from the network layer is used as the operative KBC value for ranking and extracting the upper percentile of structurally strategic segments.

4.4 Analytical Workflow

The analytical workflow consists of four main steps.

First, KBC is read on the local network, and the subset of structurally strategic segments is identified through the top 10% threshold described above. This threshold is not understood as a rigid ontological distinction between important and non-important streets. Rather, it is used as an operational cut-off that makes it possible to isolate the most clearly bridge-like segments, while still preserving a readable corridor structure.

Second, the ground-truthed premises are spatially related to the street network. Each georeferenced record is associated with its nearest segment, and a short-distance rule is then applied to determine whether it can be considered located on, or immediately adjacent to, a high-KBC corridor. The distance threshold is deliberately kept small and operational, at 12 m, because the purpose is not to model service catchments, but to identify corridor-level correspondence between structural importance and actual functional presence. The empirical distance distribution supports this choice, as it captures the great majority of clear frontage relations while excluding more ambiguous associations.

Third, the relation between premises and high-KBC segments is interpreted through the two diagnostic notions that structure the paper. Alignment refers to cases in which structurally strategic corridors already host a meaningful concentration of local functions. Mismatch refers instead to corridors that are structurally important in the KBC reading, but display weak functional activation in terms of premise presence. At this stage, the analysis also considers the mesocategories of the associated functions, not in order to produce a full typological model, but to observe whether existing and latent corridors appear to differ in the kinds of activities they currently support.

Fourth, mismatches are interpreted as latent corridors: underused structural assets whose bridging role is not yet mirrored by the present functional geography. The purpose is not to claim that such corridors should automatically become intervention sites, nor that structural prominence alone is sufficient to attract activities. Rather, the aim is to identify locations that deserve closer planning attention because they combine configurational relevance with incomplete functional activation. In this sense, the output of the workflow is not a predictive model of land-use success, but a decision-support layer for discussing where proximity-oriented actions may be more consequential within the broader EMC2 agenda (Fig. 2).

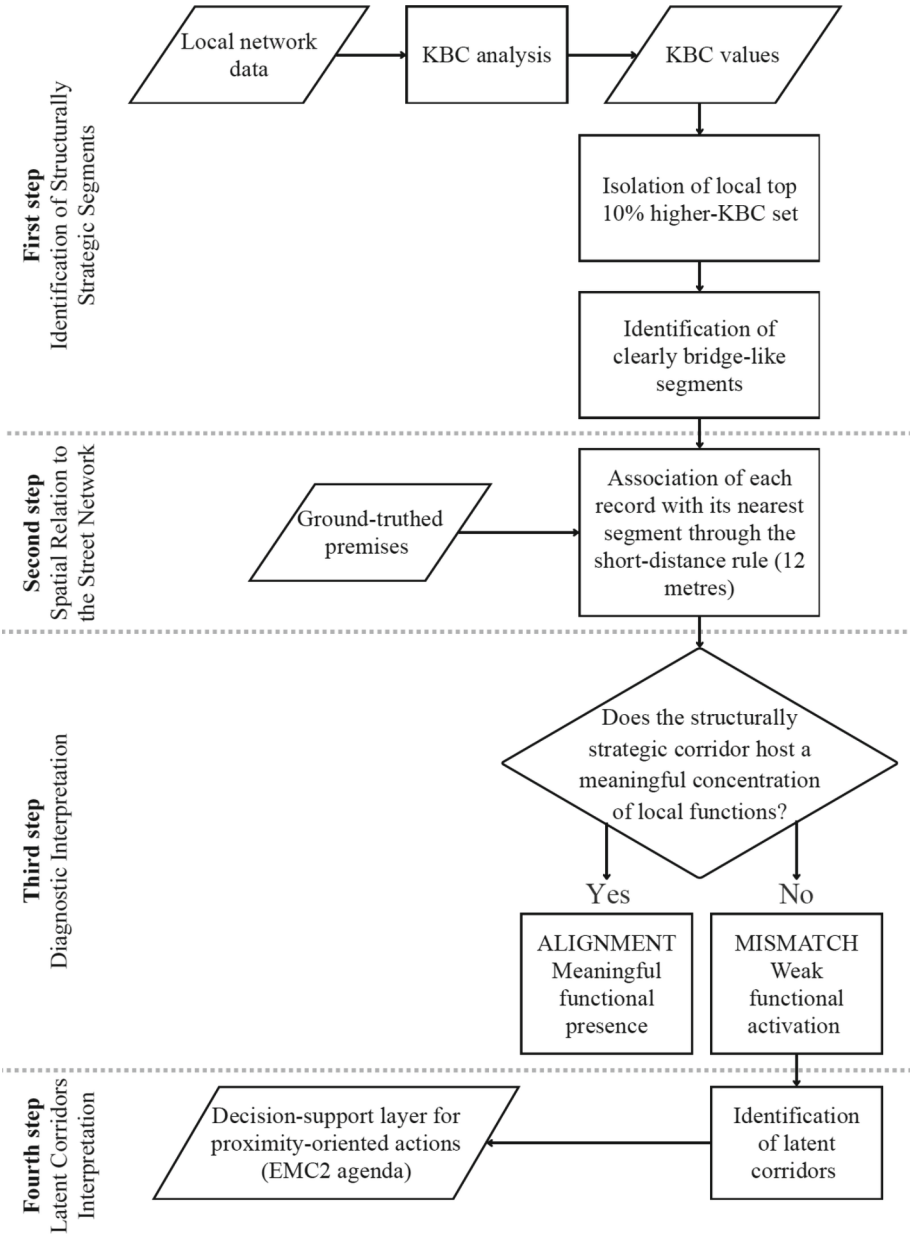


Fig. 2. Analytical workflow for relating high-KBC corridors to the ground-truthed distribution of POIs and distinguishing between existing and latent corridors.

5 Results

5.1 Bridging Structure of the Local System

The KBC analysis identifies a limited subset of segments that act as the main bridging corridors of the local system. These segments do not simply correspond to the most visible or geometrically prominent streets. Rather, they emerge because they connect otherwise differentiated parts of the network and because their removal would significantly reduce stochastic navigability across the local structure. In Turano, the high-KBC set produces a corridor pattern that strongly supports the broader project interpretation: the local main street derives from a weak tie within the Versilia system and is reinforced by a historically consolidated radial logic.

More specifically, the analysis shows that the two most evident high-KBC axes are Via Aurelia and Via delle Gorine. Via Aurelia appears as a short but highly intense bridge-like filament, expressed through three clusters of top-10% segments with markedly high values. Via delle Gorine, by contrast, appears as a longer corridor whose high-KBC structure is articulated in two disconnected top-10% stretches. These are separated by a central portion that falls below the 90th percentile, while still maintaining medium-to-high KBC values. Immediately below these two dominant axes, Via della Quercia emerges as another relevant corridor, although none of its segments enters the top-10% threshold.

Overall, the local top-10% structure includes 49 segments. These do not define a single dominant continuous street, but rather a small set of linked and hierarchised sub-corridors. This confirms that local bridgeness in Turano is selective and strongly concentrated, rather than broadly distributed across the settlement.

5.2 Current Geography of Local Functions

The ground-truthed premises layer reveals a local geography that is neither uniformly scattered nor fully corridor-based. Instead, it shows a mixed pattern, in which one axis, Via Aurelia, already concentrates the largest share of commercial and service functions, while the rest of the area remains much more fragmented. As noted above, Via Aurelia alone hosts 38 premises, including 18 active activities, and therefore already acts as the principal functional spine of the area.

Outside this axis, only small and discontinuous clusters are visible, especially around Via del Papino, Via San Remigio di Sotto, and a few isolated points along minor streets. Via delle Gorine, despite its clear bridging role in the KBC structure, currently hosts only 3 activities, all occupied, and therefore remains much more weakly activated in functional terms (Fig. 3).

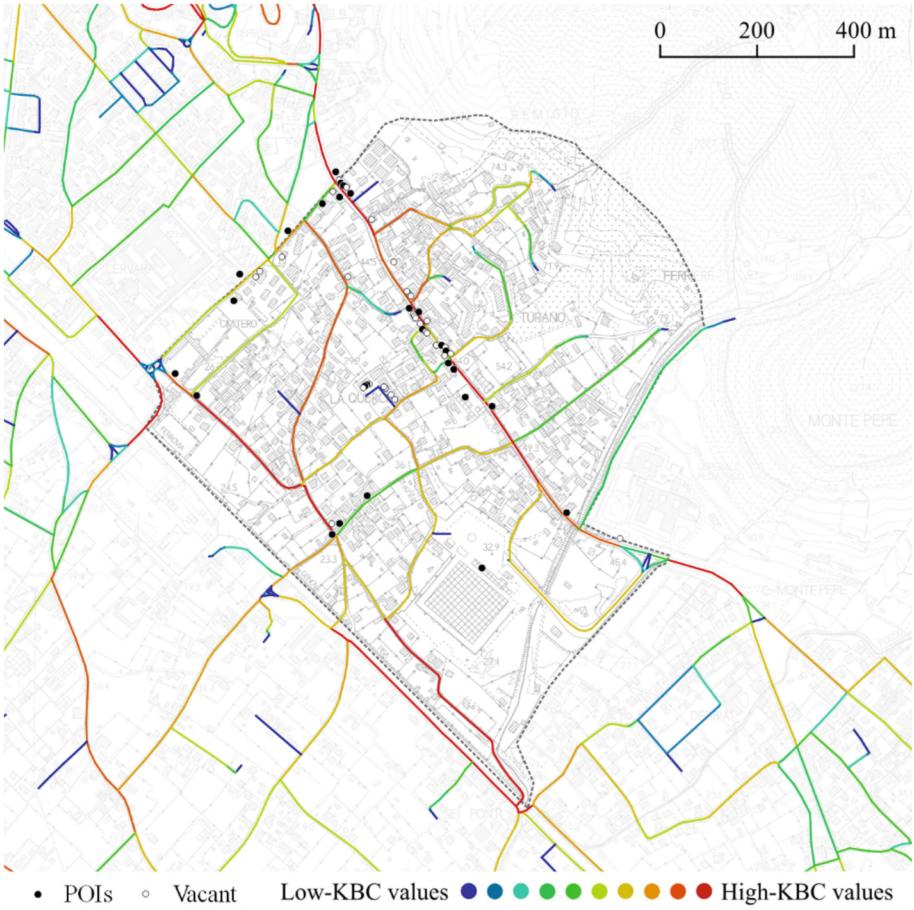


Fig. 3. KBC distribution with POIs.

5.3 Alignment and Mismatch: From Activated Corridors to Latent Assets

Going beyond visual inspection, the overlay allows a first quantitative reading of the relation between high-KBC segments and actual premises. Using nearest-segment association with a 12-m tolerance, 17 out of 63 premises are located on, or immediately adjacent to, the top 10% KBC segments. If only occupied activities are considered, the count falls to 8 out of 32. This means that the high-KBC structure captures a meaningful but still limited part of the present functional geography. In other words, the relation between bridgeness and current activity is real, but partial.

The most evident case of alignment is Via Aurelia. It is not only part of the local top-KBC structure, but also the street that already hosts the broadest and most diverse concentration of functions. In total, 14 premises fall on, or immediately beside, the Aurelia top-10% segments, of which 6 are occupied activities and 8 are vacant units. The

active uses directly associated with these top-KBC Aurelia segments include clothing and accessories, two personal-care activities, one vehicle-related activity, one pharmacy, one bakery, and one tobacco shop. Beyond the top-10% subset, the full Via Aurelia corridor hosts an even broader mix, including additional personal-care activities, tobacco and tickets, restaurant services, bars and clubs, animals and plants, furniture and home furnishings, vehicles, food and beverages, and private agencies. This confirms that Via Aurelia already operates as the most functionally activated local axis, and remains the strongest candidate to play the role of main street within the EMC2 framework.

The case of Via delle Gorine is more ambiguous, and therefore more interesting. At the street scale, it contains only 3 premises, all occupied. Of these, 2 fall directly on the top-10% KBC segments. Their typological profile is narrow, consisting of one bars-and-clubs activity and one animals-and-plants activity. A third nearby vacant unit, located on Via Bagaglione, is captured by the same high-KBC corridor. The Gorine corridor therefore shows only limited functional alignment, despite its strong structural role. Put differently, its bridging logic is much stronger than its present functional consolidation.

This is where mismatch becomes analytically useful. In Turano, mismatch does not mean that a corridor is empty or irrelevant. It means that a corridor combines high structural bridgeness with weak or narrow functional activation. Via delle Gorine is the clearest example. Its long, segmented high-KBC profile suggests a structurally strategic axis, capable of linking different parts of the local system; yet its current activity base remains extremely thin. This makes it a latent corridor in a strong sense: not a street that should automatically be transformed, but an axis whose structural importance is not yet mirrored by a comparable local-service or public-space role.

A second, weaker mismatch is visible in the minor high-KBC branch on Via del Papino, which enters the top-10% structure only marginally and currently captures no activities through the 12-m alignment rule. This residual branch is less significant than Via delle Gorine, but it confirms that high bridgeness does not automatically attract functional concentration.

Taken together, the overlay does not produce a binary answer, such as “weak ties host functions” or “they do not”. It produces a more nuanced urban reading. Via Aurelia appears as an already activated bridging corridor, where structure and functions are largely coherent. Via delle Gorine appears instead as a latent bridging corridor, where structural relevance is clear but current functional activation remains weak and narrowly specialised. This distinction is precisely what simple service-count approaches tend to miss.

6 Discussion and Conclusion

This paper set out to test a limited but important question: whether the structurally strategic corridors identified through KBC in a suburban test area already coincide with the actual geography of local functions, or whether they reveal underused assets. The answer is not framed as a claim of causality, but as a diagnostic proposition. Bridging structure alone does not generate services. However, comparing structure and function is meaningful because it helps distinguish between corridors that already work as local anchors and corridors whose structural relevance is not yet reflected in the current amenity pattern.

This contribution is useful for at least three reasons.

First, it adds a missing layer to proximity-oriented planning. Much of the current X-minute literature focuses on accessibility outputs. This is necessary, but not sufficient. A place can be assessed through thresholds, travel times, or service counts without asking whether its internal street structure contains corridors that are intrinsically more consequential than others for inter-neighbourhood linkage and local articulation. The present paper does not replace accessibility analysis. Rather, it complements it with a bridge-sensitive reading of the network.

Second, it keeps the discussion grounded in real urban conditions. The use of ground-truthed premises is not a secondary technical improvement; it is central to the validity of the argument. Earlier work on Turano and Bicchio already showed that open-data layers can produce misleading local interpretations. If the paper had relied on unreliable platform POIs, any observed alignment or mismatch could simply have reflected data noise. By using field-verified amenities, the present reading becomes more robust and more relevant for planning.

Third, the paper fits the EMC2 logic in a precise way. EMC2 does not ask only whether a suburban area currently satisfies a 15-min benchmark. It asks how suburban areas can transition towards more viable, connected, and liveable configurations organised around main streets and mixed uses. From this perspective, KBC-POI alignment is not an academic curiosity. It is one way of identifying where the local structure is already coherent with that ambition, and where it is not.

From this standpoint, the results support a differentiated planning interpretation. Via Aurelia remains the strongest candidate to act as the local main street sought by EMC2. It already combines strong bridging capacity, high visibility in the local functional geography, and the broadest diversity of current uses. It is therefore not merely a structurally strategic axis, but an already consolidated functional spine that could be reinforced through public-space improvements, better walkability, stronger cycling continuity, and selective infill or upgrading actions.

Via delle Gorine, by contrast, appears less as an already consolidated main street and more as a candidate for a secondary local centre. Its long high-KBC structure shows that it has significant intrinsic capacity to bind together the local system. Yet this capacity is currently only weakly mirrored in the geography of activities. Precisely for this reason, it seems particularly suitable for future densification and targeted activation within a broader logic of 15-min systems, where more than one local core may coexist and interact rather than depending on a single dominant centre.

The two axes should therefore not be interpreted as alternatives, but as complementary elements of a more articulated suburban strategy. In a mature EMC2 perspective, Via Aurelia would remain the principal structuring corridor, while Via delle Gorine could evolve into a supporting centre or hinge within the local system. The effectiveness of such a structure would depend less on simply adding activities and more on strategically integrating the two axes through reinforced pedestrian and cycling connections, safer and more continuous public-space interfaces, and a stronger spatial relation between everyday services and movement corridors. In this sense, the present analysis does not prescribe a finished project, but helps identify where the spatial preconditions for such a project are already present.

The paper also has clear limits. It does not claim that KBC predicts land-use success. It does not measure demand, profitability, ownership constraints, or regulatory feasibility. It does not distinguish in detail between all types of functions, although the extension towards category-specific affinity remains open. Nor does it solve the broader question of which functions should be prioritised in suburban proximity planning. On that point, the parallel DPI-related work remains essential, because different amenities follow different logics of frequency, threshold, substitutability, and capillarity. The present paper deliberately stops one step earlier. It asks where structural opportunities appear to exist, not which exact functions should necessarily occupy them.

A further limitation concerns the use of KBC itself. Recent work makes clear that KBC can be more exposed to boundary-induced artefacts than alternative removal-based measures. This is why the present analysis relies on a wider network extent and then interprets the values within the local boundary. Even with this precaution, the metric should still be treated as an exploratory planning-support layer rather than as a definitive ranking of urban importance.

Within these limits, the paper still offers a useful and original message for MLC. In suburban areas, streets should not be evaluated only by how many services they currently host or how quickly they connect to existing destinations. Some of them also deserve attention because they are structurally strategic. When such corridors are compared with the real geography of local functions, they reveal a distinction that matters for planning: already activated corridors versus latent corridors. This distinction can support more informed discussions about where to reinforce local services, improve pedestrian continuity, prioritise cycling links, upgrade bus-stop environments, or test more selective forms of densification. It does not settle those decisions, but it helps structure them.

Future work can develop this diagnostic in three directions. The first is empirical refinement, through the inclusion of category-specific analyses to test whether some amenity types display stronger affinities with high-KBC corridors than others. The second is analytical integration, by combining bridgeness with travel-time accessibility and with the more explicit functional prioritisation now being developed through DPI-related work. The third is operational extension, by linking latent corridors to ownership, regulatory, and implementation layers already present in the broader EMC2 workflow.

In conclusion, the main contribution of the paper is modest but clear. It shows that weak ties are useful for suburban planning not simply because they identify bridges in the abstract, but because, when compared with the actual distribution of local functions, they help reveal whether structurally strategic corridors are already working as local anchors or remain underused spatial assets. In Turano, this distinction leads to a particularly relevant planning implication: Via Aurelia emerges as the most plausible existing main street, while Via delle Gorine appears as a promising corridor for future activation and densification within a more integrated 15-min suburban system.

Acknowledgments. This work was part of the EMC2 project, funded by ANR (France), FFG (Austria), MUR (Italy) and Vinnova (Sweden) under the Driving Urban Transition Partnership, which has been funded by the European Commission.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Papadopoulos, E., Sdoukopoulos, A., Politis, I.: Measuring compliance with the 15-minute city concept: state-of-the-art, major components and further requirements. Sustain. Cities Soc. **96**, 104875 (2023)
2. Pezzica, C., Altafini, D., Mara, F., Chioni, C.: Unpacking proximity modelling for x-minute cities: a systematic methodological review. Sustainability **18**(9), 4469 (2026). https://doi.org/10.3390/su18094469
3. Cutini, V.: The weak ties that bind a city: 'to see bridges where others see holes'. In: Proceedings of the 14th International Space Syntax Symposium, Nicosia (2024)
4. Fusco, G., Pont, M.B., Cutini, V., Psenner, A.: Guiding principles for the 15-minute city in peripheral areas: the EMC2 model. In: Game Changer? Planning for Just and Sustainable Urban Regions, pp. 690–707 (2024)
5. Burt, R.S.: The network structure of social capital. Res. Organ. Behav. **22**, 345–423 (2000)
6. Altafini, D., Bini, D.A., Cutini, V., Meini, B., Poloni, F.: An edge centrality measure based on the Kemeny constant. SIAM J. Matrix Anal. Appl. **44**(2), 648–669 (2023)
7. Mara, F., Anselmi, C., Deri, F.: Urban Systems, When They Conurbate: The Diachronic-configurational approach to detect the shifting of centrality in Versilia. Urban Design International (2025)
8. Hillier, B.: Space is the Machine: A Configurational Theory of Architecture. Cambridge University Press, Cambridge (1996)
9. Meltzer, R., Capperis, S.: Neighbourhood differences in retail turnover: evidence from New York City. Urban Stud. **54**(13), 3022–3057 (2017)
10. Dolega, L., Lord, A.: Exploring the geography of retail success and decline: a case study of the Liverpool City Region. Cities **96**, 102456 (2020)
11. Anselmi, C., Mara, F., Deri, F.: Big-data reliability in informing the 15-min city modelling: a comparison between OpenStreetMap and Google Maps. In: Gervasi, O., et al. Computational Science and Its Applications – ICCSA 2025 Workshops. ICCSA 2025. Lecture Notes in Computer Science, vol. 15892. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97638-4_21
12. Deri, F., Mara, F., Anselmi, C.: Crowdsourced data for urban planning: a critical evaluation of OpenStreetMap accuracy and completeness. In: Gervasi, O., et al. Computational Science and Its Applications – ICCSA 2025 Workshops. ICCSA 2025. Lecture Notes in Computer Science, vol. 15897. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97660-5_27
13. Mara, F., Deri, F., Anselmi, C.: Mismapped cities: the risks of using big-data in urban planning. In: De Lotto, R., Venco, E., Pietra, C. (eds.) Innovation in Urban and Regional Planning. INPUT 2025. Lecture Notes in Civil Engineering, vol. 826. Springer, Cham (2026). https://doi.org/10.1007/978-3-032-18822-9_38

Mathematical Methods for Image Processing and Understanding (MMIPU 2026)

On Xylographic Image Reconstruction: the Segmentation Phase

Ivan Gerace(✉)

Laboratorio di Matematica Computazionale Sauro Tulipani, Dipartimento di Matematica e Informatica, Università degli Studi di Perugia, Perugia, Italy
ivan.gerace@unipg.it

Abstract. This paper addresses the reconstruction of images from xenographic documents, which often exhibit multiple and varied types of distortions. To effectively resolve these issues, a multi-stage reconstruction process is required, with each stage targeting a specific distortion. A primary challenge in this process is the segmentation of pixels according to the printer's color palette. This study proposes incorporating the segmentation phase at an early stage of reconstruction and introduces a technique to achieve this objective. The experimental results demonstrate promising outcomes for facilitating subsequent reconstruction stages.

Keywords: image reconstruction · image segmentation · color spaces · document reconstruction

Introduction

Since the 14th century, card games have become widespread in Europe, leading to the emergence of various types of local playing cards [15,16,18]. While many of these card types remain in use, others have disappeared. In certain instances, only historical decks produced using xylographic printing techniques and subsequently deteriorated over time have survived. The *Perugine* type is one such example.

A multi-stage technique for the digital reconstruction of such decks was proposed in [13]. The present work revisits and modifies this technique by introducing an additional stage: segmentation [12,22]. Detailed analysis of the reconstruction of xylographic images reveals that the core challenge lies in segmenting pixels according to the limited hues available to historical printers. Following a defading stage [1,9,19,20] that restores the chromatic components of the image to their original saturation, segmentation becomes a crucial step. The segmented image subsequently facilitates the reconstruction of borders and colored regions. Accordingly, a chromatic segmentation technique is proposed that utilizes a uniform color space [4], such as QTLC-r^*g^* (Quadratically Transformed Luminance Chrominance-r^*g^*). This color space is defined by a polynomial transformation from RGB coordinates (see [6]). Employing a uniform color space minimizes discrepancies in visual error during evaluation.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 77–94, 2027.
https://doi.org/10.1007/978-3-032-30530-5_5

The structure of this paper is as follows. Section 1 reviews the printing systems used to produce playing cards. Section 2 discusses the significance of the *Perugine* deck for reconstruction efforts. Section 3 examines the direct problem of producing historical playing cards using xylographic techniques. Section 4 presents the technique proposed in [13] for reconstructing xylographic images. Section 5 introduces the proposed technique for the new segmentation phase, and Sect. 6 presents the experimental results associated with this approach.

1 Printing Systems

Printing techniques for playing cards have varied significantly since their inception. Decks intended for the lower classes were mass-produced using matrices engraved on wooden boards, whereas tarot cards used in courts were hand-painted by renowned artists. Printing systems remained artisanal until the early nineteenth century, when technological advances enabled mass production. The following subsection briefly introduces the most commonly used printing methods [2,5].

1.1 Xylography

The first printing technique used for producing playing cards was *woodcut* or *xylography*. This method uses wooden blocks carved with chisels. The resulting lines are quite bold and clear. The design is created by the protruding part of the block, and the ink is applied to this relief. The inks, both the black for the base design and those for coloring the image, were made by dissolving dye powders in water, along with substances that thickened the color and fixed it to the paper. The colors, initially applied by hand, began to be applied using stencils, with shapes corresponding to the areas to be colored. Woodcut remained in use, at least for the less expensive cards, until the early 20th century. Unfortunately, woodcuts are characterized by often broken lines, as the wooden block has poor resistance to both weathering and wear. Colors are added with stamps or stencils, so in some places their deviation from the black lines and the colored areas that need to be filled in are noticeable. Figure 1 illustrates a wooden block utilized in this printing technique.

1.2 Chalcography

A technique called *chalcography* uses metal plates, rather than wooden ones. This technique began to be used for printing playing cards in the 17th century. Specifically, the metal plate was engraved with a burin, and ink was deposited in the resulting cavity. This technique allowed for finer lines than with woodcuts, thus achieving more realistic shading. Because the ink was in the plate's cavity, to achieve a perfect print, the pressure between the plate and the paper had to be greater than in woodcuts. This resulted in numerous raised areas on the newly printed sheet, making stencil coloring more difficult. Therefore, the paper was often colored with watercolors.

Fig. 1. wooden block for printing *Siciliani* Tarot cards.

1.3 Lithography

Toward the end of the 19th century, a printing method called *lithography* was developed. It involved using a well-polished limestone plate on which the design was traced with a grease pencil. By spreading a special substance on the wet plate, the drawn portion would hold it, while the color would not adhere to the rest of the stone. Drawing with a grease pencil was easier than producing a plate and allowed corrections. With this technique, the colors were printed first, and then the figures' outlines were traced in black. Since black has the highest covering power, this method hides any imperfections or smudges in the underlying colors. Thus, the colors appear well-aligned.

1.4 Modern Techniques

Lithography's tones are very flat, with colors lacking nuances or shading. To overcome this problem, the *chromolithography* technique was developed. This technique uses colored dots to create shadows or color gradients. The *photolithography* technique was developed in the second half of the 19th century. With this technique, printing is not done directly: the image is reproduced on a sheet of paper or a rubber plate, and then transferred to the final sheet. This system also allows a photograph to be used as a matrix. Finally, *dot lithography* uses colored dots obtained with halftones or electronically. However, not the entire image is screened, and the rest of the drawing uses solid colors.

2 Deck of Cards Under Reconstruction

While the English-style deck of cards is the most widely recognized international standard, many regions maintain distinct local decks that support the card games prevalent in their respective cultures [15,16,18].

These decks exhibit a high degree of standardization, as players seldom accept minor variations. In Italy, specifically, decks are often associated with individual regions. This regional diversity originated because, prior to 1861, Italy consisted of several independent states with distinct customs regulations, prompting each

state to develop its own local card designs. Within the former Papal States, the *Piacentine*-style deck of cards was prevalent, despite the fact that the city of Piacenza itself was geographically close but not part of the Papal States. In contrast, the classification of playing cards in the area surrounding Perugia (Umbria) remains more contentious. The following analysis examines a deck that, until the 1930 s, was identified as *Perugine*.

Guglielmo Murari was likely the most influential playing card manufacturer in Italy from the late nineteenth to the early twentieth century [10]. Born in Treviso, Veneto (then part of the Austrian Empire), in 1847, Murari was the son and grandson of established card manufacturers. The family relocated to Bari (Apulia) in 1867, where the business continued. In 1871, Guglielmo Murari assumed control of the enterprise. The Bari factory produced a significant volume of cards, manufacturing all local Italian patterns for distribution in regional markets. Among these products was the *Perugine* deck, which is illustrated in Fig. 2.

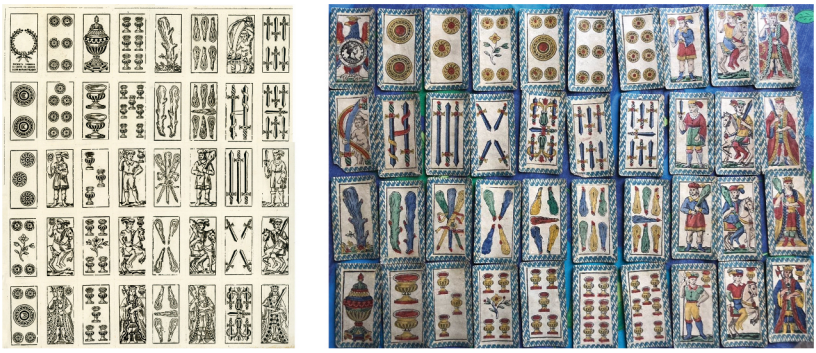
Fig. 2. *Perugine* cards produced by Murari (from Trinei's private collection).

During the second half of the nineteenth century in Perugia, the card manufacturer Verga was active. Verga produced large *Toscane* cards (see Fig. 3 (a)), small *Toscane* cards, and *Perugine* cards (see Fig. 3 (b)). The *Toscane* cards were likely intended for the nearby province of Arezzo in Tuscany, whereas the *Perugine* cards were certainly produced for local use. Subsequently, another Perugian manufacturer, Buoncompagni, produced large *Toscane* cards (see Fig. 3 (b)), while Bartolini manufactured *Perugine* cards (see Fig. 3 (c)). In the nearby city of Foligno in Umbria, Mariani also produced *Perugine*-type cards (see Fig. 4 (a)). Additionally, Mariani created a *Perugine*-type deck featuring a *Piacentine* ace of coins (see Fig. 4 (b)).

The Solesio family was among the most prominent card manufacturers. The Solesios originally produced playing cards in Spain and relocated to Genoa in the early nineteenth century. In the early twentieth century, the Solesio company manufactured *Perugine*-style decks of cards (see Fig. 5 (a)) and distributed them wrapped in sheets of paper engraved with "di tipo Perugia" (see Fig. 5 (b)).

(a) (b) (c)

Fig. 3. (a): large *Toscane* cards produced by Verga; (b): in the first row: small *Toscane* cards produced by Verga; in the second row: large *Toscane* cards produced by Buoncompagni; in the third row: *Perugine* cards produced by Verga; (c): *Perugine* cards produced by Bartolini [21].



(a) (b)

Fig. 4. (a) and (b): *Perugine* cards produced by Mariani;

(a) (b)

Fig. 5. (a): *Perugine* cards produced by Solesio; (b): playing card wrapper by the Solesio company.

The *Perugine* type of cards has been discontinued since the 1930 s, making such decks extremely rare. This article examines a digital reconstruction of the Solesio deck with the objective of producing a new reprint, thereby enabling players in the Perugia area to reconnect with their gaming heritage.

3 The Direct Problem

In this section, we formalize the direct problem of image formation by xylographic techniques. Assume we have an ideal image composed of $u + 1$, with $u \geq 1$, subimages: one containing only the color black, and the other u containing a single, distinct color.

The black image contains only black and white pixels, while the colored images contain only pixels of a specific color or white. In these images, it is assumed that for each pixel, only at most one of the four subimages has a non-white pixel, except in the case of overlapping black and colored pixels in the internal regions of the image, where black does not correspond to the edge of an object. Furthermore, we assume that, in colored images, the colors completely cover connected regions of non-black pixels. Ideally, the black image contains the borders delimiting the colored areas with a nearly uniform thickness across the entire image, while other black objects depicting details can have different thicknesses.

During the printing phase, the resulting image ideally has a black pixel if it is black in the black image; otherwise, a colored pixel if it is colored with one of the three colors, or white if it is not colored in any of the four images. During this phase, due to xylographic printing, various alterations can occur. Due to wear and tear on the wooden printing press, the black image may exhibit alterations at the edges, where black pixels may be added or removed, replaced by new black, white, or intermediate-colored pixels. Note that the number of pixels added or removed at the edge varies, potentially altering the thickness of the black border. In some places, the black border may even be broken (see Fig. 6 (a) and (b)). This fact can be modeled using an appropriate probabilistic model, which must be determined on a case-by-case basis. The size of colored images may also vary at the edges due to printing; furthermore, some connected color areas may be slightly translated and rotated, resulting in imperfect overlap with the black border (see Fig. 6 (a)-(c)) and, in some areas, the overlapping of two colors. In the latter case, the resulting image will appear as a subtractive synthesis of the two overlapping colors.

Due to the temporal deterioration of the playing cards, there is a loss of saturation in all image components, including the white areas, as well as a slight blurring and the addition of noise, which we assume to be white and independent Gaussian. The blurring and loss of saturation may not be uniform across the image.

Finally, small color patches may be added to the image , whose colors depend on the area where they are inserted (see Fig. 6 (c)). This entire model presents many probabilistic variables and will be formalized and tested for validity in a subsequent work.

(a) (b) (c)

Fig. 6. (a): detail of the two of clubs, showing discontinuous black borders and red paint leakage beyond the borders; (b): detail of the ace of spades, illustrating discontinuous borders and an incomplete green area; (c): detail of the four of spades, indicating artificial color in some areas and an incomplete blue region. (Color figure online)

4 The Xylographic Image Reconstruction Problem

According to [13], the woodcut image reconstruction technique can be divided into five distinct phases. The first phase addresses damage resulting from paper and ink deterioration, which causes the data images to appear faded [1,9,19,20]. The defading technique proposed in [6] incorporates a deblurring step to reconstruct the diffusion of the paper ink. The second phase focuses on estimating the black border to separate various color regions. This task is classified as a curve reconstruction problem (see also [3,17]), and in [13], it is solved manually using established software tools. The third phase introduces a reconstruction process to remove artifacts and enhance color homogeneity, utilizing a regularization technique [7,8,11]. The fourth phase involves detecting the outer contour, which is distorted in the data due to the xylographic printing process and must therefore be estimated. To address this, an original regularization technique is employed, based on an energy function with three components: a coherence term to maintain proximity to the initial boundary, and two additional terms to regulate the boundary's length and thickness. The fifth and final phase estimates the intensity of undefined colors, using a variant of a classic sparse data resolution technique [14].

5 Segmentation Phase

This paper introduces a novel segmentation phase [12,22], positioned between the defading phase and the phase in which the black border of the image is reconstructed to address breaks generated during printing. This segmentation phase is essential, as it facilitates more direct management of subsequent processing steps. Furthermore, the xylographic reconstruction problem aligns more closely with a segmentation problem than with deblurring or regularization challenges. It is necessary to identify non-zero pixels in both the black and color images, as

well as to determine the specific color present in each color image. Consequently, the image must be segmented into $u + 1$ subimages.

To address the segmentation problem effectively, it is necessary to select a uniform color space in which adjacent colors appear similar to the human eye [4]. Specifically, we must group pixels with similar colors and assign them non-zero values. In this paper, the chosen color space aligns with that used in [13], namely the QTLC-r^*g^* (Quadratically Transformed Luminance Chrominance-r^*g^*) space, as introduced in [6]. This space is not only uniform but also supports the application of quadratic transforms when converting to the standard RGB space. The following subsection provides further details regarding this color space.

5.1 The QTLC-r^*g^* Color Space

Given a color $(r, g, b) \in [0, 255]^3$ in the RGB space, we initialy consider

$$\begin{cases} l = (1/4)r + (1/2)g + (1/4)b, \\ r^* = \dfrac{r-l}{l}, \qquad g^* = \dfrac{g-l}{l}. \end{cases}$$

l is the luminance coordinate of the color in the QTLC-r_*g_* space, the chrominance coordinates are computed as follows

$$\begin{cases} \alpha = (5/8)r^* - (3/8)g^* + (1/4)\, r^*\, g^* \\ \beta = (3/8)r^* + (5/8)g^*. \end{cases}$$

Given the image in Fig. 7 (a), Fig. 7 (b) presents its luminance; Fig. 7 (c) shows its chrominance; the component α and β are given in Fig. 7 (d), and (e), respectively.

(a) (b) (c) (d) (e)

Fig. 7. (a): king of coins of the *Piacentine* desk produced by Dal Negro; (b): the luminance component; (c): the chrominance component with constant luminance equal to 180; (d) the α component; (e) the β component.

To segment an image based on the colors used during the printing phase, it is necessary to analyze the chromatic components rather than the luminance.

Focusing on chrominance allows for the exclusion of variations in color intensity that are represented by luminance. Figure 8 shows a circular section of radius 1 of the chromatic components, with luminance set to a value of 180.

Fig. 8. chromatic space.

Let $t = \sqrt{(34+16\beta)^2 - 24(40\alpha + 24\beta)}$, then the inverse transforms from the QTLC-r^*g^* space into the RGB space is given by

$$\begin{cases} r = \frac{l}{12}(34 + 16\beta - t) + l, \\ g = \frac{l}{20}(-34 + 16\beta + t) + l, \\ b = 4l - r - 2g. \end{cases}$$

5.2 Proposed Techninique

Let image $X^{(1)} = (L^{(1)}, A^{(1)}, B^{(1)}) \in \mathbb{R}^{n\times m\times 3}$ be the faded image obtained after the first phase, where $L^{(1)}$ is the luminance component, and $A^{(1)}$ and $B^{(1)}$ the two components α and β, respectively.

We discretize the chromatic components on a grid with a fixed step h. Let $M(\alpha, \beta)$ be the number of occurrences of the chrominance component (α, β) in the image $X^{(1)}$. Now we divide the chrominance map into 4 areas associated with high-frequency chrominances, so that each area is almost connected. We do not require each map to be connected, as small chrominance bands may be randomly missing due to missing occurrences. We proceed as follows.

Since image $X^{(1)}$ is the result of a fading algorithm, its colors are highly saturated. We can assume that the maximum value of M occurs at $\alpha = \beta = 0$, where we have colors like black and white, and possibly some shades of gray.

We begin by creating a rectangular area on the map, centered at $\alpha = \beta = 0$, in which we will evaluate which high-frequency chrominances are connected to the chrominance $(0, 0)$. Note that we have to limit this area to a rectangle otherwise, we would find all the high-occurrence areas.

Consider the following procedure.

procedure Black
for $\alpha \in \{-2, 2\}$, with step h **do**
  for $\beta \in \{-2, 2\}$, with step h **do**

$f(\alpha, \beta) = 0$;
end for
end for
for $\alpha \in \{\widetilde{\alpha}, \overline{\alpha}\}$, with step h **do**
for $\beta \in \{\widetilde{\beta}, \overline{\beta}\}$, with step h **do**
$f(\alpha, \beta) = 1$;
end for
end for
$f(0, 0) = 2$;
for $\widetilde{h} \in \{-h, h\}$ **do**
for $\overline{h} \in \{-h, h\}$ **do**
for $\alpha \in \{\widetilde{\alpha}, \overline{\alpha}\}$, with step $\widetilde{h}$ **do**
for $\beta \in \{\widetilde{\beta}, \overline{\beta}\}$, with step $\overline{h}$ **do**
if $M(\alpha, \beta) > \varepsilon \wedge d(2, \alpha, \beta) \wedge f(\alpha, \beta) = 1$ **then**
$f(\alpha, \beta) = 2$;
end if
end for
end for
end for
end for
end procedure

By the first for loops, we impose that the partition to which all the color components belong is unknown, that is, $f(\alpha, \beta) = 0$ for each pair (α, β). The second for loops imposes that the color components within a rectangle centered on the origin are selected, so $f(\alpha, \beta) = 1$ for $(\alpha, \beta) \in [\widetilde{\alpha}, \overline{\alpha}] \times [\widetilde{\beta}, \overline{\beta}]$. The origin is inserted into the gray partition, so $f(0, 0) = 2$. In subsequent for loops, the rectangle centered on the origin is scanned four times in all possible directions, thus determining all the color components of the gray partition. If: the frequency of a particular component is high enough (i.e., $M(\alpha, \beta) > \varepsilon$, with ε an appropriate threshold); the distance between (α, β) and the gray partition is less than an appropriate threshold (in this case the value of the function $d(2, \alpha, \beta)$ is true); (α, β) is not yet part of the gray partition (i.e., $f(\alpha, \beta) = 1$); then (α, β) is placed in the gray partition (i.e., $f(\alpha, \beta) = 2$).

Now let's examine the following procedure.

procedure Chromatic Segmentation
for $k \in \{3, 4, \ldots, u + 2\}$; **do**
$(\hat{\alpha}, \hat{\beta}) = \arg\max_{f(\alpha,\beta)=0} M(\alpha, \beta)$;
$f(\hat{\alpha}, \hat{\beta}) = i$;
let $Q(l)$ be the occurrences of $(l, \hat{\alpha}, \hat{\beta})$ in $X^{(1)}$;
$l^{(2)}(k) = \arg\max_{l \in [0,255]} Q(l)$;
$\alpha^{(2)}(k) = \hat{\alpha}$;
$\beta^{(2)}(k) = \hat{\beta}$;
for $\widetilde{h} \in \{-h, h\}$ **do**
for $\overline{h} \in \{-h, h\}$ **do**

```
      for α ∈ {α̃, ᾱ}, with step h̃ do
        for β ∈ {β̃, β̄}, with step h̄ do
          if M(α, β) > ε ∧ d(k, α, β) ∧ f(α, β) = 0 then
            f(α, β) = k;
          end if
        end for
      end for
    end for
  end for
end procedure
```

This procedure has u principal cycles, each of which determines a chromatic partition, denoted by $k \in \{3, 4, \ldots, u+2\}$. Initially, the maximizer $(\hat{\alpha}, \hat{\beta})$ of the occurrences of the components not yet assigned to any partition, i.e., such that $f(\alpha, \beta) = 0$, is determined. The component $(\hat{\alpha}, \hat{\beta})$ is thus assigned to the k-th partition, i.e., $f(\hat{\alpha}, \hat{\beta}) = k$. For each possible luminance l, the vector $Q(l)$ of the occurrences of $(l, \hat{\alpha}, \hat{\beta})$ in the image under consideration $X^{(1)}$ is determined. It is established that all pixels with a chromatic component belonging to the k-th partition will be assigned a luminance equal to the mode of the luminances of pixels with a chromatic component equal to $(\hat{\alpha}, \hat{\beta})$. Furthermore, all pixels of the partition under consideration will be assigned the chromatic component $(\hat{\alpha}, \hat{\beta})$. With successive for loops, the entire chromatic grid is scanned in all directions, and if: the frequency of a particular component is high enough (i.e. $M(\alpha, \beta) > \varepsilon$); the distance to the partition k is less than a suitable threshold ($d(k, \alpha, \beta)$ is true); (α, β) is not part of any partition ($f(\alpha, \beta) = 0$); then (α, β) is placed in partition k (i.e. $f(\alpha, \beta) = k$).

Finally, let's consider the main program.

```
main program
Black;
Chromatic Segmentation;
for i ∈ {1, ..., n} do
  for j ∈ {1, ..., m} do
    if f(α(i, j), β(i, j)) > 2 then
      α(i, j) = α^(2)(f(α(i, j), β(i, j)));
      β(i, j) = β^(2)(f(α(i, j), β(i, j)));
      l(i, j) = l^(2)(f(α(i, j), β(i, j)));
    end if
    if f(α(i, j), β(i, j)) = 2 then
      α(i, j) = 0;
      β(i, j) = 0;
      if l(i, j) < s then
        l(i, j) = 255;
      else
        l(i, j) = 0;
      end if
```

```
      end if
      if f(α(i,j),β(i,j)) = 1 then
        α(i,j) = 0;
        β(i,j) = 0;
        l(i,j) = 0;
      end if
      if f(α(i,j),β(i,j)) = 0 then
        l(i,j) = −1;
      end if
    end for
  end for
  while there exists (i,j) such that l(i,j) = −1 do
    for i ∈ {1,...,n} do
      for j ∈ {1,...,m} do
        if l(i,j) = −1 then
          update(l(i,j),α(i,j),β(i,j));
        end if
      end for
    end for
  end while
```

The main program begins by calling the two procedures, Black and Chromatic Segmentation. We have assumed that all variables in the two procedures are global and therefore visible in the main program. All pixels in the image $X^{(1)}$ are examined. If the partition assigned to the chromatic component of that pixel is a colored partition, i.e., $k \in \{3, 4, \ldots, u+2\}$, then the pixel's color components are assigned the values relative to the k-th partition. If the chromatic component of pixel (i, j) belongs to the gray partition, we proceed as follows: if the luminance exceeds a threshold s, the pixel is assigned the color white; otherwise, it is assigned the color black. We have decided to assign the color black to pixels whose chromatic components fall within the central rectangle but are not part of the gray partition (i.e., such that $f(\alpha(i,j),\beta(i,j)) = 1$). For pixels for which the chromatic component is indeterminate (i.e., $f(\alpha(i,j),\beta(i,j)) = 0$), we proceed as follows. We set the pixel value to indeterminate ($l(i,j) = -1$) and, as long as there are indeterminate pixels, we loop over them and update their value with the mode of the already determined pixels around them (function update$(l,(i,j),\alpha(i,j),\beta(i,j))$).

6 Experimental Results

The proposed technique was implemented in the C language on a Linux platform. Experimental results are reported for three different 1080×2000 data images. Because the original distortion-free images are not available, quantitative assessment of reconstruction quality is not feasible. Consequently, the quality of the results is evaluated visually. To this end, in this paper, the results of the proposed technique are compared with those obtained using the denoising technique

given in [8]. Our technique outcome serves as an intermediate step intended to enhance the quality of subsequent reconstructions, including automatic border closure and uniform coloring of enclosed areas.

Note that it is not possible to manually create the ideal original image, as this process would require arbitrary color selection. Consequently, any error measurement between the estimated image and the manually created image would primarily reflect differences in color choices rather than algorithmic performance. The proposed technique selects colors from the chromatic component of a uniform color space, which ensures more accurate and consistent color selection. In contrast, comparing the proposed technique with an alternative segmentation method is not feasible, as such a method would require significant simplification and adaptation to align with the specific requirements of the current problem. Additionally, fundamental differences in results would again arise from unpredictable color choices. Therefore, the proposed algorithm is compared with a regularization and denoising technique. The denoising method employed is sophisticated and incurs higher computational costs than the proposed approach; however, simpler denoising techniques may exist that achieve comparable results with reduced computational time. It is important to note that the output generated by the proposed algorithm serves as an intermediate step to facilitate

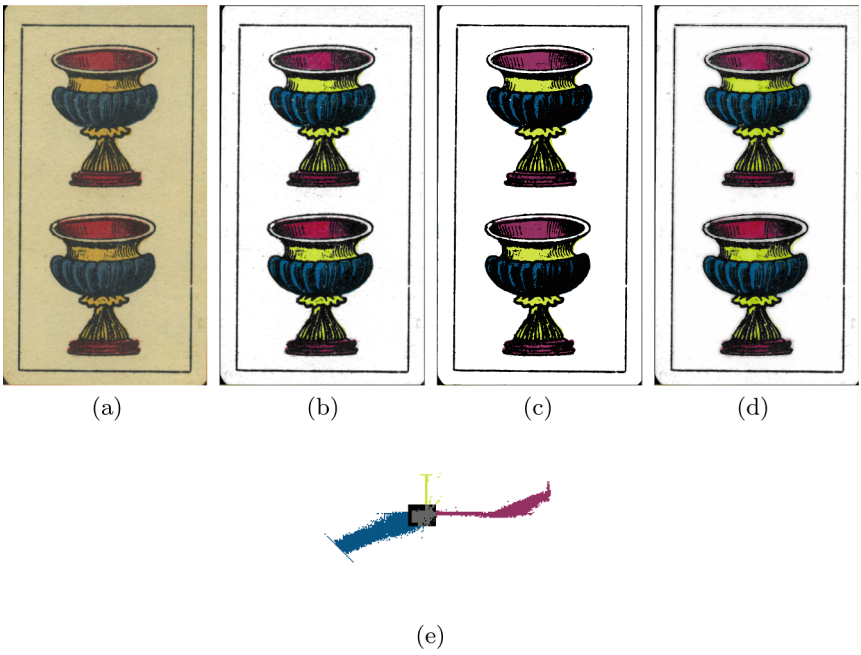
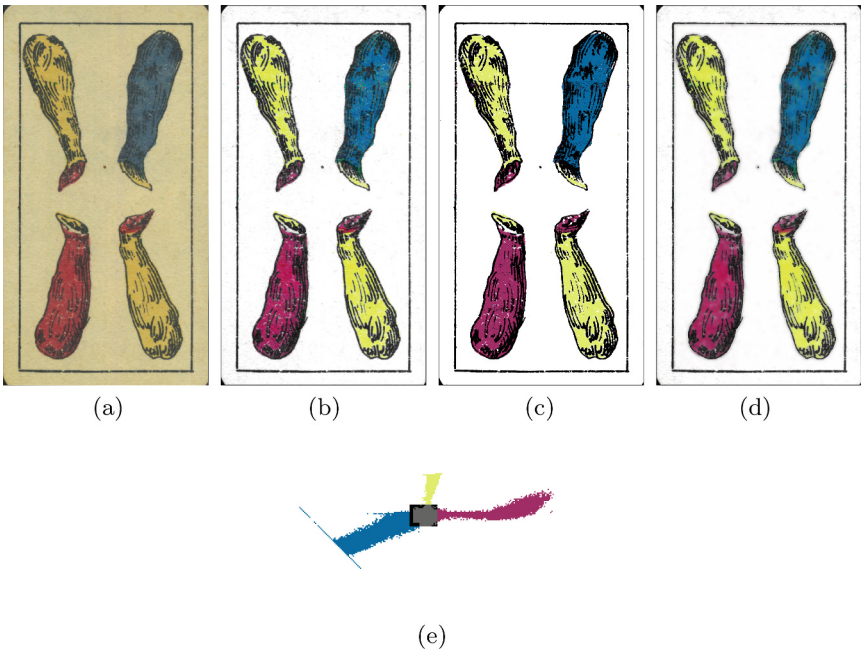


Fig. 9. (a): first data image; (b): defaded image; (c): segmented image; (d) denoised image; (e) segmented chrominance space.



(a) (b) (c) (d)

(e)

Fig. 10. (a): second data image; (b): defaded image; (c): segmented image; (d) denoised image; (e) segmented chrominance space.

the subsequent phase, which was previously performed manually and is currently under investigation.

The parameters utilized to evaluate the proposed algorithm are presented below. $h = 0.01$, $\widetilde{\alpha} = -0.15$, $\overline{\alpha} = 0.07$, $\widetilde{\beta} = -0.20$, $\overline{\beta} = 0.07$, $\varepsilon = 10$, and $s = 128$. These parameters were determined through experimental procedures.

Experimental results for the three images are shown in Figs. 9, 10, and 11. Each figure includes: (a) the data image to be processed; (b) the defaded data, denoted as $X^{(1)}$; (c) the output of the algorithm using three colors, denoted as $X^{(2)}$; (c) the reconstruction by the algorithm in [8]; and (d) the chromatic partition map given by the procedures Black and Chromatic Segmentation presented in the Subsect. 5.2. In the partition map, chromatic components classified as grey are shown in grey, components of the rectangle centered on the origin are shown in black, and colored components are displayed in their respective colors.

Figure 12 presents detailed views of the input images and the corresponding results produced by the proposed technique and the technique in [8]. In Figs. 12 (a), (d), and (e), the reconstructed image displays a completely white upper edge of the cup. In contrast, applying the denoising method yields a light gray edge. Examination of Figs. 12 (b), (e), and (f) shows that the green spots within the blue area are entirely removed, and the blue region remains unchanged. Note that the small details that remain incorrect can be easily removed in subsequent

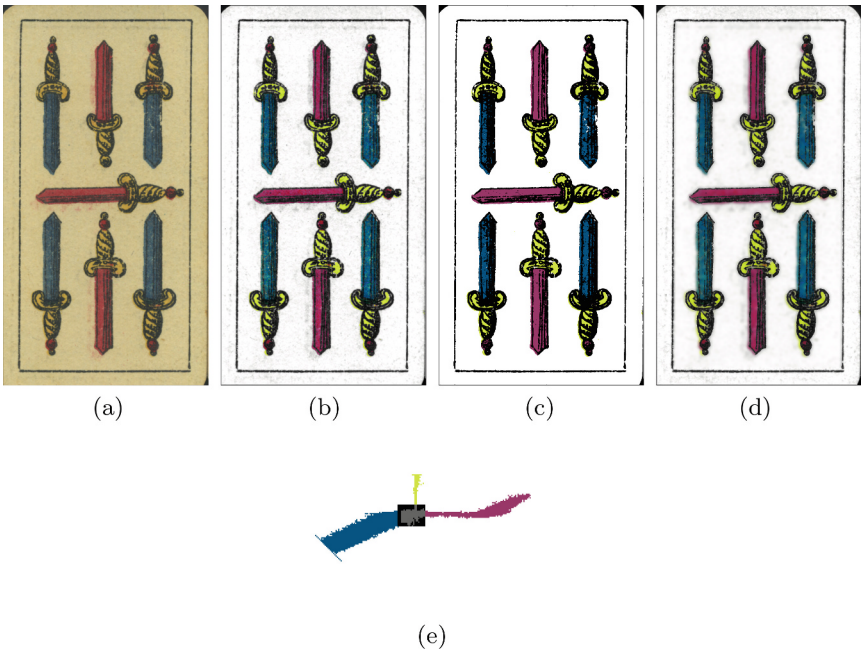


Fig. 11. (a): third data image; (b): defaded image; (c): segmented image; (d) denoised image; (e) segmented chrominance space.

phases. The denoising technique, however, alters the color. Finally, Figs. 12 (c), (f), and (i) demonstrate that the proposed technique effectively eliminates halos present in the data to be reconstructed. This result differs from that of the proposed denoising.

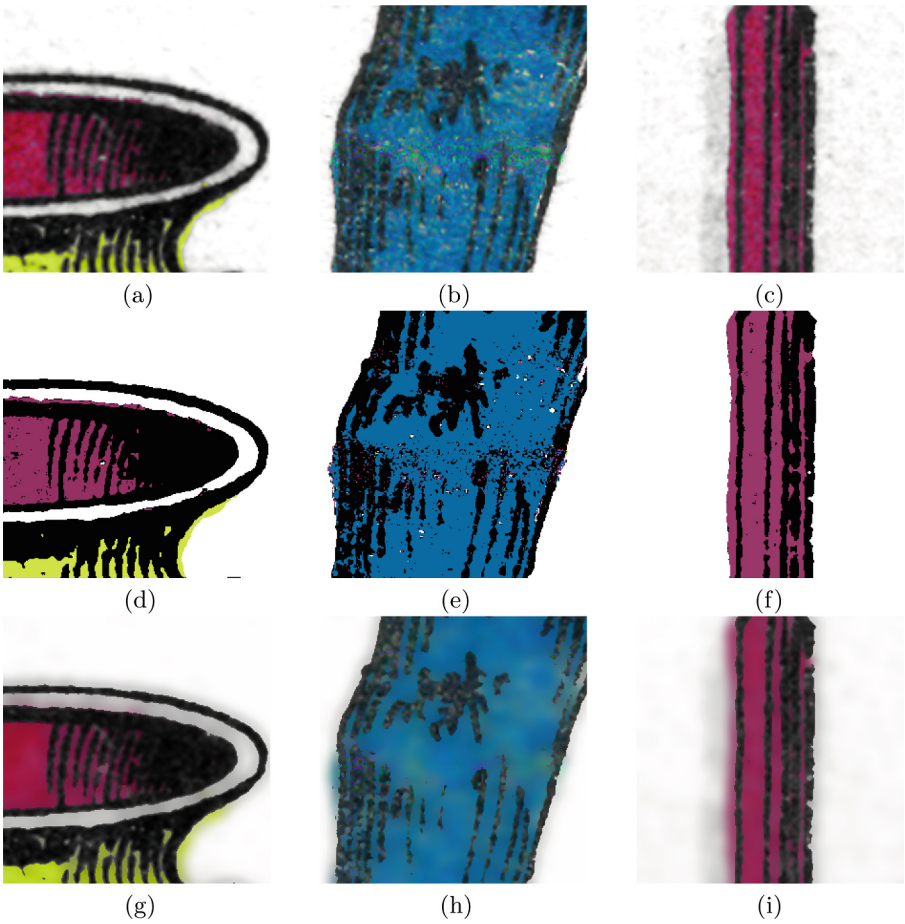


Fig. 12. (a): first defaded image detail; (b): second defaded image detail; (c): third defaded image detail; (d): first segmented image detail; (e): second segmented image detail; (f): third segmented image detail;(g): first denoised image detail; (h): second denoised image detail; (i): third denoised image detail.

Conclusions

This paper addresses the problem of reconstructing images of xylographic documents, noting that distortions in such images can manifest in various forms. Previous work [13] examined this issue in the context of reconstructing an ancient deck of *Perugine* playing cards, dividing the process into several phases. The present study introduces a new phase that segments pixels according to the printer's hues. Specifically, a technique based on the segmentation of the chromatic components within the uniform color space QTLC-r^*g^* (see [6]) is proposed. The experimental results are promising, particularly regarding their

application in subsequent reconstruction phases. A potential extension of this research involves evaluating all cards in the deck collectively to determine the overall print hues of the deck.

References

1. Ahmed, A. M. T.: Color restoration techniques for faded colors of old photos, printings, and paintings. In: Proceedings of 2009 IEEE International Conference on Electro/Information Technology, pp. 151–156 (2009)
2. Associazione 7bello – Cartagiocofilia Italiana A.P.S. https://www.7bellonline.it/. Accessed 20 Mar 2025
3. Barzegar Khalilsaraei, S., Komar, A., Zheng, J., Augsdörfer, U.: InceptCurves: curve reconstruction using an inception network. Vis. Comput. **40**, 4805–4815 (2024)
4. Basova, O., et al.: Uniform color space with advanced hue linearity: PCS23-UCS. In: Schettini, R., Trémeau, A., Tominaga, S., Bianco, S., Buzzelli, M. (eds) Computational Color Imaging. CCIW 2024. LNCS, vol 15193. Springer, Cham (2025)
5. Beal, G.: The printing of playing-cards. Playing-Card **29**(3), 106–114 (2000)
6. Biondi, G., Boccuto, A., Gerace, I.: Quadratically transformed luminance chrominance spaces. In: Gervasi, O., et al. Computational Science and Its Applications – ICCSA 2023 Workshops. ICCSA 2023. LNCS, vol. 14108. Springer, Cham. 676–693 (2023)
7. Boccuto, A., Gerace, I., Giorgetti, V.: A graduated non-convexity technique for dealing large point spread functions. Appl. Sci. **2** (10), 28 (2023)
8. Boccuto, A., Gerace, I., Martinelli, F.: Half-quadratic image restoration with a non-parallelism constraint. J. Math. Imaging Vision **59**(2), 270–295 (2017)
9. Bruni, V., Ramponi, G., Restrepo, A., Vitulano, D.: Context-based defading of archive photographs. EURASIP J. Image Video Process. **2009**(986183), 11 (2009)
10. De Giorgio N. A.: La manifattura delle carte da gioco di Guglielmo Murari. Comune di Veglie (2005)
11. Demoment, G.: Image reconstruction and restoration: overview of common estimation structures and problems. IEEE Trans. Acoust., Speech Signal Process. **37**, 2024–2036 (1989)
12. Erb, W.: Split-and-merge segmentation of biomedical images using graph wedgelet decompositions. In: Gervasi, O., et al. Computational Science and Its Applications – ICCSA 2025 Workshops. ICCSA 2025. LNCS, vol 15899. Springer, Cham. 252–263 (2025)
13. Gerace, I.: Digital reconstruction of ancient paper artifacts. In: Gervasi, O., et al. Computational Science and Its Applications – ICCSA 2025 Workshops. ICCSA 2025. LNCS, vol 15892. Springer, Cham, pp. 400–416 (2025)
14. Gerace, I., Palomba, C., Tonazzini, A.: An inpainting technique based on regularization to remove bleed-through from ancient documents. In: Proceedings of the 2016 International Workshop on Computational Intelligence for Multimedia Understanding (IWCIM), Reggio Calabria, Italy, pp. 1–5 (2016)
15. Hargrave, C.P.: An History of Playing Cards. Courier Corporation, United States (2014)
16. Jilley, R.: A History of Playing cards. Studio Vista, London (1973)
17. Laville, B., Blanc-Féraud, L., Aubert, G.: A Γ-convergence result and an off-the-grid charge algorithm for curve reconstruction in inverse problems. J. Math. Imaging Vision **66**, 572–583 (2024)

18. Mann, S.: Collecting Playing Cards. Howard Baker Press (1966)
19. Nikitenko, D., Wirth, M., Trudel, K.: Applicability of white-balancing algorithms to restoring faded colour slides: an empirical evaluation. J. Multimed. **3**(5), 9–18 (2008)
20. Restrepo, A., Ramponi, G.: Filtering and luminance correction for aged photographs. In: Image Processing: Algorithms and Systems VI, Proceedings of SPIE, vol. 6812, pp. 26–31, San Jose, Calif, USA (2008)
21. Tarot-as-Tarocchi. https://www.tarot-as-tarocchi.com/files/piacentine-perugia--1-.jpg. Accessed 17 Mar 2025
22. Zaitoun, N.M., Aqe, M.J.: Survey on image segmentation techniques. Proc. Comput. Sci. **65**, 797–806 (2015)

The Minimum Sum Coloring Problem on Circulant Graphs $C_n(1, 2)$

Eugenio Maria Cocco[1] and Sara Nicoloso[2(✉)]

[1] Università di Roma Tor Vergata, Via del Politecnico 1, 00133 Roma, Italy
[2] Istituto di Analisi dei Sistemi ed Informatica, CNR, Via dei Taurini 19, 00185 Roma, Italy
sara.nicoloso@iasi.cnr.it

Abstract. A circulant graph $C_n(a, b)$ is a graph with n vertices $\{v_0, \ldots, v_{n-1}\}$ such that each vertex v_i is adjacent to vertices $v_{(i+a) \bmod n}$ and $v_{(i+b) \bmod n}$. The Minimum Sum Coloring problem consists in finding a vertex coloring $\mathcal{C} : V \to \mathbb{Z}_+$ in such a way that adjacent vertices receive a different color and $\sum_{v_i \in V} c_i$ is minimum. In this paper, we investigate the Minimum Sum Coloring problem on the connected circulant graphs $C_n(a, b)$ with $n \geq 6$, $\gcd(n, a) = 1$, and $b = (\pm 2a) \bmod n$. After a brief discussion about isomorphism, using an array representation of the graphs, we show that an optimal Odd Cycle Transversal can be computed in linear time and in closed form on these graphs, as well as the independence number, the minimum Chromatic Sum, and an optimal Sum Coloring solution. The approach we propose shows that the Sum Coloring problem on these graphs can be solved to optimality in $O(n)$ time and that the strength $s(\cdot)$ of the graph, i.e. the number of colors needed to achieve the minimum Sum Coloring, equals the chromatic number $\chi(\cdot)$ of the graph.

Keywords: sum coloring · circulant graphs · independent set

1 Introduction

Consider $t + 1$ non-negative distinct integers $n, a_1, a_2, ..., a_t$ such that $n > 0$, $a_i \in \{1, \ldots, n-1\}$ for $i = 1, .., t$, and verifying $a_i + a_j \neq n$ for $i, j = 1, .., t$, $i \neq j$. An (undirected) *circulant graph* $C_n(a_1, a_2, ..., a_t) = (V, E)$ has $V = \{v_0, v_1, \ldots, v_{n-1}\}$ as vertex set, and $E = \{(v_i, v_{(i+a_h) \bmod n})$ for $i = 0, ..., n-1$ and $h = 1, ..., t\}$ as edge set [1,2]. Since $\big((i + a_h) \bmod n + (n - a_h)\big) \bmod n = i$, for an arbitrary $i = 0, ..., n-1$ and for an arbitrary entry a_h, the edges $(v_i, v_{(i+a_h) \bmod n})$ and $(v_{(i+a_h) \bmod n}, v_{((i+a_h) \bmod n+(n-a_h)) \bmod n})$ coincide. Thus, without loss of generality, we can limit ourselves to consider $a_i \in \{1, \ldots, \lfloor \frac{n}{2} \rfloor\}$ for $i = 1, ..., t$. $C_n(a_1, a_2, ..., a_t)$ is $2t$-regular if $a_i \neq \frac{n}{2}$ for $i = 1, .., t$ and $(2t-1)$-regular otherwise.

We will consider (simple undirected) connected circulant graphs with $t = 2$ entries, denoted by a and b, for the sake of simplicity. The graph will be denoted

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 95–107, 2027.
https://doi.org/10.1007/978-3-032-30530-5_6

as $C_n(a, b)$ (see Fig. 1). The hypotheses on a and b, also called *entries*, imply that $n \geq 4$. Notice that the graph $C_n(a, b)$ consists of $p = \gcd(n, a, b)$ copies of (the connected) $C_{\frac{n}{p}}\left(\frac{a}{p}, \frac{b}{p}\right)$. Hence $C_n(a, b)$ is connected if and only if $\gcd(n, a, b) = 1$ [3].

It is easy to see that, given n, a, and b, $C_n(a', b')$ and $C_n(b', a')$ with $a' \in \{a, n - a\}$ and $b' \in \{b, n - b\}$ are all the same graph.

Circulant graphs form a class of highly symmetric mathematical structures [4], are a special case of the Cayley graph, and are a natural extension of a ring with increased connectivity. These graphs are used to model several practical applications, for example, in the field of small-world networks, discrete cellular neural networks, supercomputing, chemical reactions, and multiprocessor systems [5].

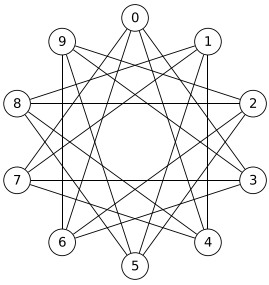


Fig. 1. The 4-regular circulant graph $C_{10}(3, 4)$.

Given a simple (i.e. without loops) graph $G = (V, E)$, a *vertex colouring* $\mathcal{C} : V \to \mathbb{Z}_+$ is a function which assigns a non–negative integer value to each vertex of the graph. A vertex colouring $\mathcal{C}$ is *feasible* if adjacent vertices receive different colors. The *Vertex Colouring Problem* (VCP, for short) consists in finding a feasible vertex colouring which uses the minimum number, $\chi(G)$, of colors. $\chi(G)$ is called the *chromatic number* of G.

The chromatic number of circulant graphs $C_n(a, b)$ is fully characterized by the following theorem.

Theorem 1. [3,6,7] *Let $C_n(a, b)$ be a connected circulant graph. Then*

$$\chi(C_n(a,b)) = \begin{cases} 2 \text{ if } n \text{ even, and } a, b \text{ odd} \\ 4 \text{ if } n = 13 \text{ and } \left(b \equiv \pm 5a \pmod{13} \text{ or } a \equiv \pm 5b \pmod{13}\right) \\ 4 \text{ if } n \neq 5, \ 3 \not| n, \text{ and } \left(b \equiv \pm 2a \pmod{n} \text{ or } a \equiv \pm 2b \pmod{n}\right) \\ 5 \text{ if } n = 5 \\ 3 \text{ otherwise.} \end{cases}$$

A slightly different problem is the *Minimum Sum Colouring Problem* (MSCP, for short). It consists in finding a feasible vertex colouring with minimum cost, where the *cost* of a feasible vertex colouring $\mathcal{C}$ for G is defined as $z(\mathcal{C}) = \sum_{v_i \in V} c_i$. Such minimum cost is the *chromatic sum* of G, and is denoted $\Sigma(G)$ (see Fig. 2).

It is also important to determine the number of colors necessary to achieve $\Sigma(G)$. This number is called the *strength* $s(G)$ of the graph. Clearly, $s(G) \geq \chi(G)$. MSCP can be generalized to the *Optimum Cost Chromatic Partition* (OCCP) problem [8], where a non decreasing sequence $(k_1, k_2, ..., k_{|V|})$ of color costs is given and the problem consists in finding a feasible vertex colouring such that $\sum_{v_i \in V} k_{c_i}$ is minimized. OCCP reduces to MSCP if $(k_1, k_2, ..., k_{|V|}) = (1, 2, ..., |V|)$.

Fig. 2. The label assigned to each vertex indicates its colour (light grey=1, light yellow=2, and green=3). The vertex colouring on the left (respectively, right) uses two (respectively, three) colours its cost is 12 (respectively, 11). Consequently, the colouring on the left is optimal for VPC as the graph is bipartite, but it is not optimal for MSCP; conversely, the colouring on the left is not optimal for VPC, but it is optimal for MSCP (thus, $\Sigma(G) = 11$ and $s(G) = 3$. (Color figure online)

Minimum Sum Colouring problems arise in many applications, for example, in VLSI design, distributed resource allocation, and scheduling [8,9].

MSCP is an NP-hard problem on general graphs [10], and, indeed, in the literature, many papers focus on computing lower and upper bounds or on heuristic and meta-heuristic algorithms, or exact algorithms for small instances or classes of graphs with special properties (see for example [8,9,11–14]).

MSCP on circulant graphs is studied in [15]. In their paper, the authors state the following proposition, where C_n denotes a cycle with n vertices and $\Sigma(C_n) = \frac{3}{2}n$ if n is even and $\Sigma(C_n) = 3 + \frac{3}{2}(n-1)$ if n is odd.

Proposition 1. *[15] Let $C_n(1,t)$ be a circulant graph, where $n = m(t+1)+r$ and $0 \leq r \leq t$. There exists a proper coloring for $C_n(1,t)$ for which the chromatic sum is*

$$\Sigma(C_n(1,t)) = \begin{cases} \Sigma(C_n) \text{ if } t \text{ is odd and } n \text{ is even} \\ (\leq)\Sigma(C_n) + 2t - 2 \text{ if } t \text{ and } n \text{ are odd} \\ m\Sigma(C_{t+1}) \text{ if } t \text{ is even and } r = 0 \\ (\leq)m\Sigma(C_{t+1}) + \frac{7}{2}r + 2 \text{ if } t \text{ and } r \text{ are even} \\ (\leq)(m-1)\Sigma(C_{t+1}) + \frac{3}{2}(r+1) + \frac{7}{2}t \text{ if } t \text{ is even and } r \text{ is odd.} \end{cases}$$

In the same paper similar results are stated for the circulant graphs $C_n(1, 2, ..., t)$.

An *independent set* (also known as *stable set*) of a graph $G = (V, E)$ is a set of mutually non-adjacent vertices. The cardinality of a largest independent set in G is the *independence number* of G and is denoted by $\alpha(G)$ [16], and the problem of determining $\alpha(G)$ for a graph G is known as the *Maximum Independent Set* problem. Observe that, if $C_n(a,b)$ is not connected $\alpha\big(C_n(a,b)\big) = p{\cdot}\alpha\big(C_{\frac{n}{p}}(\frac{a}{p}, \frac{b}{p})\big)$.

By definition, any vertex colouring $\mathcal{C} : V \to \mathbb{Z}_+$ induces a partition $\mathcal{P} = \{V_1, V_2, ..., V_k\}$ of V into $k \leq n$ independent sets, as vertices assigned the same color are non-adjacent. The viceversa is also true: every partition $\mathcal{P} = \{V_1, V_2, ..., V_k\}$ of V induces a colouring $\mathcal{C}$ where the color c_i assigned to an arbitrary vertex v_i is j if $v_i \in V_j$. For this reason, in the sequel, by vertex colouring $\mathcal{C}$ we mean a partition of V into independent sets. It is easy to see that the cost $z(\mathcal{C}) = \sum_{j=0}^{n-1} c_i$ of a vertex colouring $\mathcal{C}$ can be computed also as $z(\mathcal{C}) = \sum_{j=1}^{k} j|V_j|$. This way of computing $z(\mathcal{C})$ shows that a colouring $\mathcal{C} = \{V_1, V_2, ..., V_k\}$ whose subsets are ordered by non-increasing cardinality is one which minimizes the cost among the $k!$ equivalent vertex colouring obtained by ordering the subsets in a different way. Hence a colouring achieving the minimum chromatic sum $\Sigma(G)$ can be searched among those verifying $|V_1| \geq |V_2| \geq ... \geq |V_k|$. In the sequel we will consider colouring verifying these inequalities, only.

By what above, an immediate lower bound for the chromatic sum $\Sigma(G)$ of an arbitrary graph $G = (V, E)$ can be computed as

$$\sum_{j=1}^{\lfloor \frac{n}{\alpha(G)} \rfloor} j \cdot \alpha(G) + \left\lceil \frac{n}{\alpha(G)} \right\rceil \cdot n \bmod \alpha(G). \tag{1}$$

The proposed lower bound is computed assuming that there exists a colouring $\mathcal{C} = \{V_1, V_2, ..., V_{\lceil \frac{n}{\alpha(G)} \rceil}\}$ where all the subsets are maximal independent set, except, possibly, the last one whose cardinality is $n \bmod \alpha(G)$. Since $0 \leq n \bmod \alpha(G) \leq \alpha(G) - 1\}$, the thesis follows. Observe that if n is a multiple of $\alpha(G)$ in the last term of (1) one has $n \bmod \alpha(G) = 0$. Other lower bounds can be found in [9,11,12].

An important issue about circulant graphs is isomorphism. Two graphs $G = (V, E)$ and $G' = (V', E')$ are isomorphic if there exists a function $f : V \to V'$ such that $(u, v) \in E$ implies $\big(f(u), f(v)\big) \in E'$ [2,4,17].

A special type of isomorphism can be defined for circulant graphs. Precisely, two circulant graphs $C_n(a_1, a_2..., a_t)$ and $C_n(a'_1, a'_2, ..., a'_t)$ are said to be *Ádám-isomorphic* if there exists a positive integer $\mu \in \{1, ..., n-1\}$, coprime to n, i.e. satisfying $gcd(\mu, n) = 1$, such that $\{a'_1, a'_2, ..., a'_t\} = \{\pm\mu a_1, \pm\mu a_2, ..., \pm\mu a_t\}$, where all the operations are intended to be mod n. The integer μ is called *Ádám's multiplier*. The *Ádám's conjecture* states that two circulant graphs are isomorphic if and only if they are Ádám-isomorphic. The conjecture does not hold in general, and several counterexamples of isomorphic circulant graphs that are not Ádám-isomorphic have been found [2,4]. However, it holds for circulant graphs $C_n(a, b)$ with two entries [18–20]. Bézout Theorem states that, given n and a, there exists an integer h such that $(ha) \bmod n = \gcd(n, a)$. Therefore, it is always possible to transform an arbitrary $C_n(a, b)$ into the Ádám-isomorphic $C_n(\gcd(n, a), (hb) \bmod n)$.

In this paper we focus on MSCP on a connected circulant graph $G = C_n(a, b)$ where $n \geq 6$, $\gcd(n, a) = 1$ and $b = (\pm 2a) \bmod n$, and we characterize its strenght $s(G)$. Actually, the results about the isomorphism of a circulant graph with two entries a and b, and, in particular when $b = (\pm 2a) \bmod n$, state that

every $C_n(a, b = (\pm 2a) \bmod n)$ admits $C_n(1,2)$ as an isomorphic graph. For this reason, we will restrict ourselves to consider the circulant graphs $C_n(1,2)$. The only graphs $C_n(1,2)$ excluded by the assumption $n \geq 6$ are $C_n(1,2)$ for $n = 3,4,5$. Since $C_n(1,2) = K_n$ for $n = 3,4,5$, MSCP is trivially solved in these cases.

The paper is organized as follows. A matrix representation is described in Sect. 2. In Sect. 3 the main results. Concluding remarks are found in Sect. 4.

2 Matrix Representation

Let $C_n(a,b) = (V,E)$ be a connected circulant graph. If $(x \pm a) \bmod n = y$, then $v_x, v_y \in V$ are *a-adjacent* and (v_x, v_y) is an *a-edge*. In particular: if $y = (x+a) \bmod n$ $\quad (y = (x-a) \bmod n$, respectively) and we move from v_x to v_y on the a-edge (v_x, v_y), we say that the a-edge (v_x, v_y) is walked along in the positive (negative, resp.) direction; similarly, for the b-edges. By a-cycle we denote a simple (i.e. loopless) cycle made of a-edges only; a similar definition holds for b-cycles. By definition the graph contains $\gcd(n,a)$ distinct a-cycles with $\frac{n}{\gcd(n,a)}$ vertices each, and $\gcd(n,b)$ distinct b-cycles with $\frac{n}{\gcd(n,b)}$ vertices each.

The circulant graphs we deal with can be represented by an array with an infinite number of rows and columns, which suitably organizes the vertices and the edges of $C_n(a,b)$: the connectedness condition $\gcd(n,a,b) = 1$ ensures that every vertex of the graph is represented in $M_n^\infty(a,b)$. While each element of $M_n^\infty(a,b)$ corresponds to a vertex of the graph, each vertex is represented by infinitely many elements in $M_n^\infty(a,b)$. If $m_{i,j}^\infty$ corresponds to the vertex v_x, we shall write $m_{i,j}^\infty = x$. If $m_{i,j}^\infty = x$ then $m_{i,j-1}^\infty = (x-a) \bmod n$, $m_{i,j+1}^\infty = (x+a) \bmod n$, $m_{i-1,j}^\infty = (x-b) \bmod n$, and $m_{i+1,j}^\infty = (x+b) \bmod n$ (see Fig. 3). Similar structures are defined in [3,21,22]. By construction, the adjacency of two elements in a row (column, resp.) represents an a-edge (a b-edge, respectively) and the two elements are a-adjacent (b-adjacent, respectively). The positive direction for the a-edges is rightwards, and the negative one is leftwards. The positive direction for the b-edges is downwards, and the negative one is upwards. Notice that traversing an arbitrary row (column, resp.) of $M_n^\infty(a,b)$ from left to right (top to bottom, resp.) corresponds to infinitely cycling over the corresponding a-cycle (b-cycle, resp.) in the positive direction, and viceversa. Since an a-cycle and a b-cycle have $\frac{n}{\gcd(n,a)}$ and $\frac{n}{\gcd(n,b)}$ vertices, respectively, $m_{i,j}^\infty = m_{i,j+n/\gcd(n,a)}^\infty = m_{i+n/\gcd(n,b),j}^\infty$, for all $i,j \in \{-\infty, ..., +\infty\}$.

Without loss of generality we can choose an arbitrary element of $M_n^\infty(a,b)$ corresponding to v_0 and say that it lays in row 0 and column 0, that is to say, $m_{0,0}^\infty = 0$ (see light yellow row and column in Fig. 3). According to the positive directions of a- and b-edges, we have that positively indexed columns are found on the right of column 0, and positively indexed rows are found below row 0. As a consequence, $m_{i,j}^\infty = (ib + ja) \bmod n$.

Let $m_{p,q}^\infty = x$. By construction, the vertices represented in row p are the vertices x verifying $x \bmod \gcd(n,a) = (pb) \bmod \gcd(n,a)$. Since $(x + b\gcd(n,a)) \bmod \gcd(n,a) = x \bmod \gcd(n,a)$, the copies of vertex v_x are found in

row $p+t\gcd(n,a)$, where t denotes an arbitrary integer, but, generally speaking, in a different column.

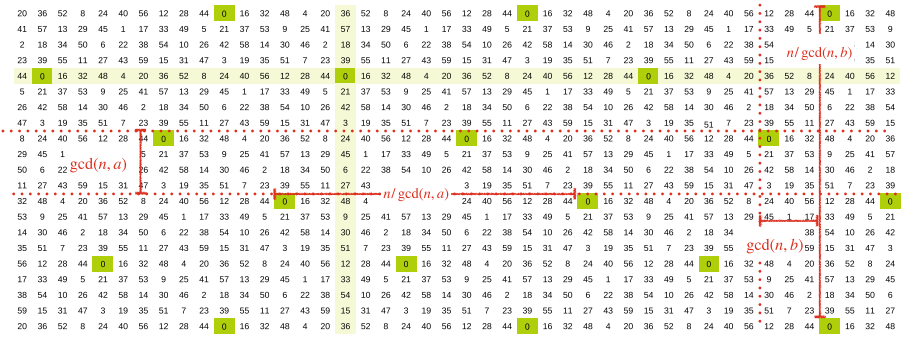


Fig. 3. Part of $M_{60}^{\infty}(16,21)$, in light yellow, row 0 and column 0; all the elements corresponding to v_0 are highlighted in green. In the example, in red, $\gcd(n,a)=4$ rows, $n/\gcd(n,a)=15$ columns, $\gcd(n,b)=3$ columns, and $n/\gcd(n,a)=20$ rows.

For the circulant graphs we deal with, the infinite array $M_n^{\infty}(1,2)$ is shown in Fig. 5 for $n=16,17$. It is easy to see that $\lambda=2$, and that if $m_{i,j}=x$ then $m_{i,j+n}=x$ and $m_{i+\gcd(n,a),j-\lambda}=m_{i+1,j-2}=x$.

3 Optimal Solutions to MSCP on Circulant Graphs $C_n(1,2)$

An Odd Cycle Transversal (OCT, for short) of a graph $G=(V,E)$ is a subset of vertices of the graph that has a nonempty intersection with every odd cycle in G. What graph theorists are interested in is finding an OCT with minimum cardinality for G, say OCT*. By definition, the graph induced by $V\backslash\text{OCT(G)}$ is bipartite, and $\text{OCT}(G)$ is empty if G is bipartite. Recalling that $m_{0,0}^{\infty}=0$, we can give the following definition.

Definition 1. [23] *Consider a circulant graph $C_n(a,b)$ and let $\mathcal{P}$ a subset of pairs of row and column indices (r,s) verifying the following conditions:*

- $m_{r,s}^{\infty}=0$,
- $(r,s)\neq(0,0)$,
- $r+s$ *even*,
- $\nexists z\in\mathbb{Z}$ *such that* $(r,s)=z\cdot(r',s')$ *where* $r'+s'$ *odd.*

Then $(r^*,s^*)\in\mathcal{P}$ is a pair of row and column indices verifying

$$\max\{|r^*|,|s^*|\}=\min\big\{\max\{|r|,|s|\}:(r,s)\in\mathcal{P}\big\}.$$

The elements $m^{\infty}_{r,s} = 0$ with $r+s$ even (odd, respectively) will be called *even (odd,* respectively) *zeroes*. Thus, in $\mathcal{P}$ there are all the even zeroes which are not even multiple of an odd zero.

Theorem 2. [23] *Let $G = C_n(a,b)$ be a circulant graph, then* $|\text{OCT}^*(G)| = \max\{|r^*|, |s^*|\}$.

Theorem 3. [23] *Let $G = C_n(a,b)$ be a circulant graph, and let* $\text{OCT}^*(G)$ *be an* OCT *with minimum cardinality for G. Then* $\alpha(G) = \frac{n-|OCT^*(G)|}{2}$.

The next theorem shows that r^* and s^*, hence $|\text{OCT}^*|$, can be computed in a closed form for the circulant graphs $G = C_n(1,2)$. In Corollary 1 we express $\alpha(C_n(1,2))$ in closed form and $|\text{OCT}^*|$ as function of it. Finally, in Theorem 5, according to [23], we build OCT^*, an OCT with minimum cardinality. We prove that either OCT^* it is an independent set itself, or it can be partitioned into two independent sets, with cardinality $\alpha(G)$ and cardinality $n \bmod \alpha(G)$, respectively. This allow us to define an optimal solution to MSCP on the $C_n(1,2)$'s, that is a vertex colouring achieving the minimum chromatic sum $\Sigma(G)$. In addition, we prove that $s(G) = \chi(G)$ colors.

Theorem 4. *Consider a circulant graph $C_n(1,2)$, with $n \geq 6$. Then*

$$(r^*, s^*) = \left(\left\lfloor \tfrac{n}{3} \right\rfloor + n \bmod 3, \left\lfloor \tfrac{n}{3} \right\rfloor - n \bmod 3\right) \quad \textit{and} \quad |\text{OCT}^*| = \left\lfloor \tfrac{n}{3} \right\rfloor + n \bmod 3.$$

Proof. The zeroes in $M^{\infty}_n(1,2)$ are found at coordinates $(t, wn-2t)$, for t, w arbitrary integers (see Fig. 4). We now show that some of them do not yield the (r^*, s^*) we are looking for.

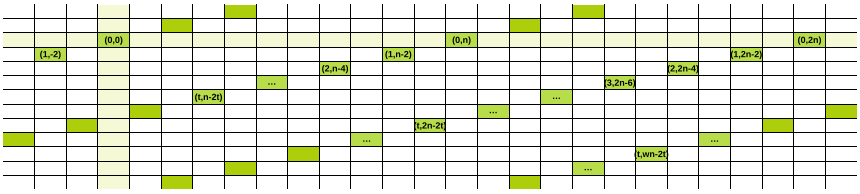


Fig. 4. The coordinates of the zeroes in $M^{\infty}_n(1,2)$

1) Recalling Definition 1, the zeroes with $w = 0$ do not belong to $\mathcal{P}$ as they are either odd zeroes or even zeroes (even) multiple of the odd one $(1,-2)$.
2) By symmetry, we can limit ourselves to consider the zeroes $(t, wn-2t)$ with $w \geq 1$, as $m^{\infty}_{t,wn-2t} = 0$ implies $m^{\infty}_{-t,-wn+2t} = 0$.

Hence, from now on we assume $w \geq 1$.

3) If wn is even, the even zero in $m^{\infty}_{0,wn}$ dominates the zeroes in $m^{\infty}_{t,wn-2t}$ for any integer $t \leq -1$, as $\max\{0, wn\} = wn < |wn-2t| = \max\{|t|, |wn-2t|\}$. If wn is odd, the even zero in $m^{\infty}_{1,wn-2}$ dominates the zeroes in $m^{\infty}_{t,wn-2t}$ for any integer $t \leq -1$, as $\max\{1, wn-2\} = wn-2 < |wn-2t| = \max\{|t|, |wn-2t|\}$.
4) For similar reasons, if wn is even, the even zeroes in $m^{\infty}_{wn/2,0}$ dominates the zero in $m^{\infty}_{t,wn-2t}$ for any integer $t \geq wn/2+1$, as $\max\{wn/2, 0\} = wn/2 < wn/2+1 \leq \max\{t, |wn-2t|\}$. If wn is odd, the even zeroes in $m^{\infty}_{\lfloor wn/2 \rfloor,1}$ dominates the zero in $m^{\infty}_{t,wn-2t}$ for any integer $t \geq \lceil wn/2 \rceil$, as $\max\{\lfloor wn/2 \rfloor, 1\} = \lfloor wn/2 \rfloor < \lceil wn/2 \rceil < \max\{|t|, |wn-2t|\}$.

Hence, from now on we will consider the zeroes in elements $m^{\infty}_{t,wn-2t}$ with $t \geq 0$ and $wn - 2t \geq 0$, only.

5) Consider an even zero Z in $(t, n-2t)$. If $w \geq 2$, Z dominates all the zeroes which are in the region South-East of it, namely all the zeroes in $(s, wn-2s)$ with $s \geq t$, $wn-2s \geq n-2t$, as $\max\{t, n-2t\} \leq \max\{s, wn-2s\}$.

Thanks to the discussion above, we search for (r^*, s^*) among the zeroes in $(t, n-2t)$ for $0 \leq t \leq \lfloor \frac{n}{2} \rfloor$ and integer, and, as observed in [23], t and $n-2t$ must have the same parity of n. By definition, the row index of the zeroes $(t, n-2t)$ increases with t and the column index $|n-2t|$ decreases with t. Hence, the maximum among row and column index is the column index until $t = n-2t$, that is for $0 \leq t \leq \frac{n}{3}$, while it is the row index for $\frac{n}{3} \leq t \leq \frac{n}{2}$. As a consequence, we can limit ourselves to consider the maximum t integer and with the desired parity in $\{0, 1, ..., \lfloor \frac{n}{3} \rfloor - 1, \lfloor \frac{n}{3} \rfloor\}$, and the minimum t integer and with the desired parity in $\{\lceil \frac{n}{3} \rceil, \lceil \frac{n}{3} \rceil + 1, ... \lfloor \frac{n}{2} \rfloor\}$,

We can limit ourselves to study the parity of n with respect to the parity of t, as $n-2t$ has always the same parity of n. Considering all these constraints, we distinguish three cases: *i*) $n \bmod 3 = 0$; *ii*) $n \bmod 3 = 1$; and *iii*) $n \bmod 3 = 2$.

Consider case $n \bmod 3 = 0$, first. Since $\frac{n}{3}$ and n have the same parity, then

$$|\text{OCT}^*| = \min\left\{ \max\{t, n-2t\} \text{ for } t \geq 0 \text{ and integer} \right\} = \frac{n}{3},$$

and $(r^*, s^*) = (\frac{n}{3}, \frac{n}{3})$, and the theses follow.

Now consider case $n \bmod 3 = 1$. Write n as $n = 3 \cdot \lfloor \frac{n}{3} \rfloor + 1$. Hence, $\lfloor \frac{n}{3} \rfloor$ and n have opposite parity. Since t and n must have the same parity, t can assume either the integer value $\lfloor \frac{n}{3} \rfloor - 1$ or the integer value $\lceil \frac{n}{3} \rceil$. We get

$$\begin{aligned}
|\text{OCT}^*| &= \min\left\{ \max\{t, n-2t\} \text{ for } t = \lfloor \tfrac{n}{3} \rfloor - 1 \text{ or } t = \lceil \tfrac{n}{3} \rceil \right\} = \\
&= \min\left\{ \max\{\lfloor \tfrac{n}{3} \rfloor - 1, n - 2(\lfloor \tfrac{n}{3} \rfloor - 1)\}, \max\{\lceil \tfrac{n}{3} \rceil, n - 2\lceil \tfrac{n}{3} \rceil\} \right. = \\
&= \min\left\{ n - 2(\lfloor \tfrac{n}{3} \rfloor - 1), \lceil \tfrac{n}{3} \rceil \right\} = \\
&= \min\left\{ \lfloor \tfrac{n}{3} \rfloor + n \bmod 3 + 2, \lceil \tfrac{n}{3} \rceil \right\} = \\
&= \lceil \tfrac{n}{3} \rceil .
\end{aligned}$$

The zero which gives such minimum is $(t, n-2t)$ for $t = \lceil \frac{n}{3} \rceil$. Hence, $(r^*, s^*) = \left(\lceil \frac{n}{3} \rceil, n - 2\lceil \frac{n}{3} \rceil \right) = \left(\lfloor \frac{n}{3} \rfloor + 1, \lfloor \frac{n}{3} \rfloor - 1 \right)$. Recalling that $n \bmod 3 = 1$, the theses follow.

The third case is when $n \bmod 3 = 2$. Write n as $n = 3 \cdot \lfloor \frac{n}{3} \rfloor + 2$. Hence, $\lfloor \frac{n}{3} \rfloor$ and n have the same parity. Since t and n must have the same parity, t can assume either the integer value $\lfloor \frac{n}{3} \rfloor$ or the integer value $\lfloor \frac{n}{3} \rfloor + 2$. We get

$$\begin{aligned} |\text{OCT}^*| &= \min\left\{ \max\{t, n-2t\} \text{ for } t = \lfloor \tfrac{n}{3} \rfloor \text{ or } t = \lfloor \tfrac{n}{3} \rfloor + 2 \right\} = \\ &= \min\left\{ \max\{\lfloor \tfrac{n}{3} \rfloor, n - 2\lfloor \tfrac{n}{3} \rfloor\}, \max\{\lfloor \tfrac{n}{3} \rfloor + 2, n - 2(\lfloor \tfrac{n}{3} \rfloor + 2)\} = \right. \\ &= \min\left\{ \lfloor \tfrac{n}{3} \rfloor + n \bmod 3, \lfloor \tfrac{n}{3} \rfloor + 2 \right\} = \\ &= \lfloor \tfrac{n}{3} \rfloor + 2. \end{aligned}$$

Since $n \bmod 3 = 2$, both zeroes considered give us the minimum OCT. That is $(r^*, s^*) \in \left\{(\lfloor \frac{n}{3} \rfloor, n - 2\lfloor \frac{n}{3} \rfloor\}), (\lfloor \frac{n}{3} \rfloor + 2, n - 2(\lfloor \frac{n}{3} \rfloor + 2))\right\} = \left\{(\lfloor \frac{n}{3} \rfloor, \lfloor \frac{n}{3} \rfloor + n \bmod 3), (\lfloor \frac{n}{3} \rfloor + n \bmod 3, \lfloor \frac{n}{3} \rfloor + n \bmod 3 - 4)\right\}$. For the sake of simplicity, we will consider as (r^*, s^*) the second one, which can be written as $(\lfloor \frac{n}{3} \rfloor + n \bmod 3, \lfloor \frac{n}{3} \rfloor - n \bmod 3)$, as $n \bmod 3 = 2$, and the theses follow.

We get the following corollary to Theorems 2, 3, 4.

Corollary 1. *Consider a circulant graph $G = C_n(1,2)$. Then*

$$\alpha(C_n(1,2)) = \lfloor \tfrac{n}{3} \rfloor \qquad \textit{and} \quad |\text{OCT}^*| = \alpha(C_n(1,2)) + n \bmod 3.$$

We can now prove the main theorem.

Theorem 5. *Consider a circulant $C_n(1,2)$. Let V_1, V_2, V_3, and V_4 be defined as follows.*

$$\begin{cases} V_1 = \{v_{1+3i}, i = 0, ..., \lfloor \frac{n}{3} \rfloor - 1\} \\ V_2 = \{v_{2+3i}, i = 0, ..., \lfloor \frac{n}{3} \rfloor - 1\} \\ V_3 = \{v_{3i}, i = 0, ..., \lfloor \frac{n}{3} \rfloor - 1\} \\ V_4 = \emptyset \end{cases} \quad \textit{if } n \bmod 3 = 0$$

$$\begin{cases} V_1 = \{v_1, v_{5+3i}, i = 0, ..., \lfloor \frac{n}{3} \rfloor - 2\} \\ V_2 = \{v_{3i}, i = 0, ..., \lfloor \frac{n}{3} \rfloor - 1\} \\ V_3 = \{v_0, v_{4+3i}, i = 0, ..., \lfloor \frac{n}{3} \rfloor - 2\} \\ V_4 = \{v_2\} \end{cases} \quad \textit{if } n \bmod 3 = 1$$

$$\begin{cases} V_1 = \{v_1, v_5, v_{9+3i}, i = 0, ..., \lfloor \frac{n}{3} \rfloor - 3\} \\ V_2 = \{v_3, v_{7+3i}, i = 0, ..., \lfloor \frac{n}{3} \rfloor - 2\} \\ V_3 = \begin{cases} \{v_0, v_4\} & \textit{if } n = 8 \\ \{v_0, v_4, v_{8+3i}, i = 0, ..., \lfloor \frac{n}{3} \rfloor - 3\} & \textit{if } n > 8 \end{cases} \\ V_4 = \{v_2, v_6\} \end{cases} \quad \textit{if } n \bmod 3 = 2.$$

Then $\mathcal{C}^* = \{V_1, V_2, V_3, V_4\}$ is an optimal solution to MSCP*,*

$$\sum(C_n(1,2)) = 2(n + n \bmod 3), \qquad \textit{and} \qquad s(C_n(1,2)) = \chi(C_n(1,2)).$$

Proof. Let $G = (V, E) = C_n(1,2)$. The proof of Theorem 5.1 in [23] states that the graph $G(V \backslash \text{OCT}^*)$ is bipartite and the two subsets of vertices have the same

cardinality $\alpha(G)$. These two subsets of vertices will be V_1 and V_2 in the vertex colouring $\mathcal{C}^*$.

According to [23], given (r^*, s^*) an optimum OCT* is obtained as we now describe.

When $n \bmod 3 = 0$, $(r^*, s^*) = (\frac{n}{3}, \frac{n}{3})$. We define OCT* as follows: $\{v_{3i}, i = 0, ..., \lfloor \frac{n}{3} \rfloor - 1\}$. We now show that this set is an independent set. In OCT* there are the vertices v_h with $h \bmod 3 = 0$, for $h = 0, ..., n-1$. The four vertices adjacent to the vertices v_h are the vertices v_k with $k = (h \pm 1) \bmod n$ or $k = (h \pm 2) \bmod n$. Since $k \bmod 3 \neq 0$, for all k, none of them belongs to OCT*. Let $V_3 =$ OCT*. Since $V_1 \cup V_2 \cup V_3 = V$ then $V_4 = \emptyset$. Therefore $\{V_1, V_2, V_3, V_4\}$ is a feasible vertex colouring. Since $|V_j| = \alpha(G)$ for $j = 1, 2, 3$, and its cost $z(\{V_1, V_2, V_3, V_4\}) = (1 + 2 + 3)\frac{n}{3} = 2n$ equals the lower bound (1), $\mathcal{C}^* = \{V_1, V_2, V_3, V_4\}$ is an optimal colouring for MSCP, and, by Theorem 1, $s(G) = \chi(G) = 3$.

When $n \bmod 3 = 1$, we define OCT* as follows: OCT$^* = \{v_0, v_2, v_{4+3i}, i = 0, ..., \lfloor \frac{n}{3} \rfloor - 2\}$ (see Fig. 5 (*a*)). This set is not an independent set as v_2 is *b*-adjacent to v_0 and v_4 (on the other hand, it could not have been an independent set as its cardinality $\alpha(G) + 1$ exceeds the cardinality $\alpha(G)$ of a maximum independent set in G). However, we can partition OCT* into $V_3 =$ OCT$^* \setminus \{v_2\}$ and $V_4 = \{v_2\}$. V_4 is trivially an independent set. As for V_3, we now prove that it is an independent set. Consider the vertices in $V_3 \setminus \{v_0\}$. These are the vertices v_h with $h \bmod 3 = 1$, for $h = 4, ..., n-1$. The four vertices adjacent to v_h are the vertices v_k with $k = (h \pm 1) \bmod n$ or $k = (h \pm 2) \bmod n$. Since $k \bmod 3 \neq 1$, for all k, none of them belongs to $V_3 \setminus \{v_0\}$. Now consider v_0. Its four adjacent vertices are v_1, v_2, v_{n-2}, and v_{n-1}. None of them belongs to $V_3 \setminus \{v_0\}$ as the indices 1 and 2 are not ≥ 4, $(n-1) \bmod 3 = 0$, and $(n-2) \bmod 3 = 2$. Hence, V_3 is an independent set, as claimed. Therefore $\{V_1, V_2, V_3, V_4\}$ is a feasible vertex colouring. Since $|V_j| = \alpha(G)$ for $j = 1, 2, 3$, $|V_4| = n \bmod \alpha(G)$, and the cost $z(\{V_1, V_2, V_3, V_4\}) = (1 + 2 + 3)\frac{n-1}{3} + 4 = 2n + 2$ equals the lower bound (1), $\mathcal{C}^* = \{V_1, V_2, V_3, V_4\}$ is an optimal colouring for MSCP, and, by Theorem 1, $s(G) = \chi(G) = 4$.

When $n \bmod 3 = 2$, we define OCT* as follows: OCT$^* = \{v_0, v_2, v_4, v_6, v_{8+3i}, i = 0, ..., \lfloor \frac{n}{3} \rfloor - 3\}$ for $n > 8$ (see Fig. 5 (*b*)) and OCT$^* = \{v_0, v_2, \ v_4, v_6\}$ for $n = 8$. This set is not an independent set as v_2 is *b*-adjacent to v_0 and v_4, and v_4 is *b*-adjacent to v_2 and v_6 (on the other hand it could not have been an independent set as its cardinality $\alpha(G) + 1$ exceeds the cardinality $\alpha(G)$ of a maximum independent set in G). However, we can partition OCT* into $V_3 =$ OCT$^* \setminus \{v_2, v_6\}$ and $V_4 = \{v_2, v_6\}$. V_4 is trivially an independent set. As for V_3, we now prove that it is an independent set. Consider the vertices in $V_3 \setminus \{v_0, v_4\}$. These are the vertices v_h with $h \bmod 3 = 2$, for $h = 8, ..., n-1$. The four vertices adjacent to v_h are the vertices v_k with $k = (h \pm 1) \bmod n$ or $k = (h \pm 2) \bmod n$. Since $k \bmod 3 \neq 2$, for all k, none of them belongs to $V_3 \setminus \{v_0, v_4\}$. Now consider v_0 and v_4. Their four+four adjacent vertices are $v_1, v_2, v_{n-2}, v_{n-1}$ and v_2, v_3, v_5, v_6. None of them belongs to $V_3 \setminus \{v_0, v_4\}$ as the indices $1, 2, 3, 5$, and 6 are not ≥ 8, $(n-1) \bmod 3 = 1$, and $(n-2) \bmod 3 = 0$.

Hence, V_3 is an independent set, as claimed. Therefore $\{V_1, V_2, V_3, V_4\}$ is a feasible vertex colouring. Since $|V_j| = \alpha(G)$ for $j = 1, 2, 3$, $|V_4| = n \bmod \alpha(G)$, and the cost $z(\{V_1, V_2, V_3, V_4\}) = (1 + 2 + 3)\frac{n-2}{3} + 4 \cdot 2 = 2n + 4$ equals the lower bound (1), $\mathcal{C}^* = \{V_1, V_2, V_3, V_4\}$ is an optimal colouring for MSCP, and, by Theorem 1, $s(G) = \chi(G) = 4$.

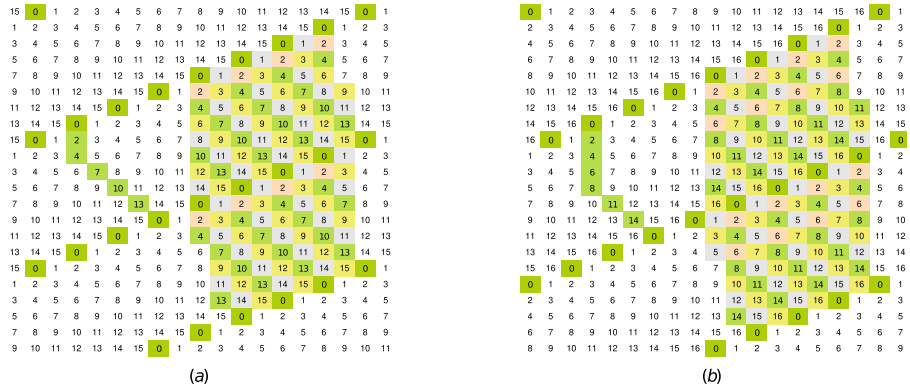


Fig. 5. (*a*) Part of $M_{16}^{\infty}(1,2)$: on the left, in green, OCT*, with 6 elements, $(r^*, s^*) = (6, 4)$; (*b*) Part of $M_{17}^{\infty}(1,2)$: on the left, in green, OCT*, with 7 elements, $(r^*, s^*) = (7, 3)$. (*a,b*) on the right, the optimum solution $\{V_1, V_2, V_3, V_4\}$ to MSCP (V_1 in light grey, V_2 in light yellow, V_3 in green, V_4 in light orange.

4 Conclusions

In this paper we focussed on the MSCP on circulant graphs $C_n(1,2)$ with $n \geq 6$. The approach is purely combinatorial and the main result is Theorem 5, where we define a vertex colouring attaining the chromatic sum $\Sigma(\cdot)$ on the circulant graphs $C_n(1,2)$ for all values of n. In addition, we show that such a coloring uses a $s(\cdot) = \chi(\cdot)$ colors.

Referring to Proposition 1, one has $m = \lfloor \frac{n}{t+1} \rfloor$, $r = n \bmod (t+1)$. When $t = 2$, in particular, one has $m = \lfloor \frac{n}{3} \rfloor$, $r = n \bmod 3 \in \{0, 1, 2\}$. Recalling that $\lfloor \frac{n}{3} \rfloor = \frac{n - n \bmod 3}{3}$, and since $\Sigma(C_3) = 6$, the results become

$$\Sigma(C_n(1,2)) = \begin{cases} 2n \text{ if } n \bmod 3 = 0 \\ (\leq) 6 \left\lfloor \frac{n}{3} \right\rfloor + \frac{7}{2} 2 + 2 = 2n + 5 \text{ if } n \bmod 3 = 2 \\ (\leq) \left(\left\lfloor \frac{n}{3} \right\rfloor - 1 \right) 6 + \frac{3}{2} \cdot 2 + \frac{7}{2} \cdot 2 = 2n + 2 \text{ if } n \bmod 3 = 1. \end{cases}$$

When $n \bmod 3 = 0$ or 1 this result coincides with our result. When $n \bmod 3 = 2$ our result states that $\Sigma(C_n(1,2)) = 2n + 4$, thus proving that the result by [15] is not tight, as the "$(\leq)$" suggests.

As a side result, we were able to express in a closed form the coordinates of an even zero corresponding to the minimum Odd Cycle Transversal of a circulant graph $C_n(1,2)$, as well as the independence number $\alpha(C_n(1,2))$.

Recalling the isomorphism issue, the results obtained in this paper apply to every $C_n(a, b = (\pm 2a) \bmod n)$ and show that MSCP on circulant graphs $C_n(a, (\pm 2a) \bmod n)$ is solvable in linear time, improving upon the general NP-hard case.

References

1. Boesch, F., Tindell, R.: Circulants and their connectivities. J. Graph Theory **8**, 487–499 (1984)
2. Elspas, B., Turner, J.: Graphs with circulant adjacency matrices. J. Comb. Theory **9**, 297–307 (1970)
3. Heuberger, C.: On planarity and colorability of circulant graphs. Discret. Math. **268**, 153–169 (2003)
4. Kamalappan, V.: A study on isomorphic properties of circulant graphs. Adv. Sci. Technol. Eng. Syst. J. **2**, 236–241 (2017)
5. El-Mesady, A.I., Romanov, A.Y., Amerikanov, A.A., Ivannikov, A.D.: On bipartite circulant graph decompositions based on cartesian and tensor products with novel topologies and deadlock-free routing. Algorithms **16** (2023). https://doi.org/10.3390/a16010010.
6. Göbel, F., Neutel, N.: Cyclic graphs. Discret. Appl. Math. **99**, 3–12 (2000)
7. Yeh, H., Zhu, X.: 4-colorable 6-regular toroidal graphs. Discret. Math. **273**, 261–274 (2003)
8. Kubicka, E.: The chromatic sum of a graph: history and recent developments. Int. J. Math. Math. Sci. **30**, 1563–1573 (2004). https://doi.org/10.1155/S0161171204306216
9. Gondran, A., Duchamp, V., Moalic, L.: Lower bound for (sum) coloring problem, manuscript, arXiv:1909.08906 (2019)
10. Kubicka, E., Schwenk, A.: An introduction to chromatic sums. In: Proceedings of the ACM Computer Science Conference, Louisville (Kentucky), pp. 39–45 (1989)
11. Lecat, C., Lucet, C., Li, C.-M.: New lower bound for the minimum sum coloring problem. In: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17), pp. 853–859 (2017)
12. Alishahi, M., Taherkhani, A.: A note on chromatic sum, manuscript. arXiv:0802.1936v2 (2009)
13. Salavatipour, M.R.: On sum coloring of graphs. Discret. Appl. Math. **127**, 477–488 (2003)
14. Nicoloso, S., Sarrafzadeh, M., Song, X.: On the sum coloring problem on interval graphs. Algorithmica **23**, 109–126 (1999)
15. Effantin, B., Kheddouci, H.: Sum Coloring of distance and circulant graphs. Electron. Notes Discret. Math. **22**, 239–244 (2005). https://doi.org/10.1016/j.endm.2005.06.058
16. Bondy, J.A., Murty, U.S.R.: Graph Theory. Springer, London (2008). https://doi.org/10.1007/978-3-662-70107-2
17. Kamalappan, V., Peraprakash, W.: Families of Type-2 isomorphic Circulant Graphs of Order np^3 *w.r.t.* $r = p$ *and their abelian groups*, manuscript, arXiv:2211.06970v1

18. Fang, X., Xu, M.: On isomorphisms of Cayley graphs of small valency. Algebra Colloq. **1**, 67–76 (1994)
19. Mans, B., Pappalardi, F., Shparlinski, I.: On the spectral Ádám property for circulant graphs. Discret. Math. **254**, 309–329 (2002)
20. Toida, S.: A note on Ádám's conjecture. J. Comb. Theory B **23**, 239–246 (1977)
21. Fiol, M.A., Yebra, L.A., Alegre, I., Valero, M.: A discrete optimization problem in local networks and data alignment. IEEE Trans. Comput. C-36, 702–713 (1987)
22. Nicoloso, S., Pietropaoli, U.: Vertex-coloring of 3-chromatic circulant graphs. Discret. Appl. Math. **229**, 121–138 (2017). https://doi.org/10.1016/j.dam.2017.05.013
23. Carotenuto, P., Nicoloso, S., Salvatore, A.: The maximum independent set problem on circulant graphs $C_n(a,b)$. Discret. Appl. Math. **384**, 411–423 (2026)

A Smoothing Aggregation Cascadic Multilevel Method for Image Deblurring

Pietro Maurino(✉) and Marco Donatelli

Department of Science and High Technology (DiSAT), Università degli Studi dell'Insubria, Como 22100, Italy
{pmaurino,marco.donatelli}@uninsubria.it

Abstract. We propose an Aggregation Cascadic Multiresolution method for image deblurring problems. Multigrid and multiresolution methods have already been studied in this context; see [4,6,7,12,13]. Here, we focus on a two-level cascadic multiresolution method with a specific choice of grid transfer operators.

The main novelty of our approach lies in the construction of both the prolongation and restriction operators, which are derived from a segmentation of the observed image. The restriction operator preserves the key features of the image by relying on an aggregation-based strategy. To further enhance the robustness of the method, a smoothing procedure is applied to the prolongation operator.

For the sake of simplicity, we employ a gradient descent method as the smoother; however, more sophisticated iterative methods can also be considered. Numerical results confirm an acceleration of convergence, even for aggressive coarsening. Moreover, in addition to accelerating convergence, a slight improvement in reconstruction quality is observed when the method is combined with a discrepancy principle stopping criterion.

Keywords: Image deblurring · Multilevel methods · Segmentation

1 Introduction

Image deblurring is a fundamental task in applied science, consisting of recovering an image from observed data that has been corrupted by measurement error or from insufficient information. Typically, an image deblurring problem is expressed as a linear system where the matrix A, which is the discretization of a convolution operator, acts on an image x, producing a blurred image b by the equation

$$Ax = b, \tag{1}$$

see [10]. The error-free image b is not available in practice due to the presence of measurement errors or due to discretization error, and we have access only to the corrupted signal b^δ such that

$$\|b - b^\delta\| < \delta, \tag{2}$$

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 108–123, 2027.
https://doi.org/10.1007/978-3-032-30530-5_7

where $\delta > 0$ is an estimation of the noise level [8]. The goal is to recover an approximation of x from an observed blurred and noisy image b^δ. In order to simulate real-world conditions, we assume a white Gaussian noise of level δ. Thus, our problem can be modeled as the linear system

$$Ax = b^\delta. \tag{3}$$

Due to the presence of noise, it is natural to look for the solution of a least square problem

$$\min_{x \in \mathbb{R}^n} \|Ax - b^\delta\|. \tag{4}$$

Since the gradient of the objective function is $A^T(b^\delta - Ax)$, where A^T is the transpose of A, the gradient descent method, also known as Landweber method [8], can be formulated as the following

$$x_{k+1} = x_k - \alpha_k A^T(Ax_k - b^\delta), \tag{5}$$

where $\alpha_k > 0$ is the step length. In practice, the typical behavior of the error norm of the computed approximation is that it decreases during the initial iterations, when the regularization error dominates, and then starts to increase after a certain iteration, as the noise error begins to amplify the norm of the reconstruction. This behavior of iterative methods applied to discretized inverse ill-posed problem is known as semi-convergence. Therefore, early stopping is crucial for obtaining accurate reconstructions. A classical strategy is to stop the iterative method (5) at the first iteration for which the norm of the residual is close to the noise level δ in (2).

Image segmentation algorithms partition an image's pixels into regions that are similar with respect to certain features. Common features include spatial proximity and grayscale similarity. Image segmentation aims to provide a lower-dimensional representation of the image, in which the identified regions offer a simpler yet effective description of the original image. The application of these algorithms in computer vision and pattern recognition has been widely successful. Various strategies have been proposed to obtain meaningful segmentations and to apply them to real-world problems [1,9].

Multigrid algorithms, originally developed for solving linear systems arising from partial differential equations, have also been applied to ill-posed problems such as image deblurring [4,6,7,12,13]. These methods are based on the definition of grid transfer operators that "transport" information across different levels of the problem in a hierarchical manner. Classical choices for prolongation and restriction operators include downsampling—typically combined with full weighting over pixel neighborhoods—interpolation, and aggregation [14,16]. Since aggregation-based grid transfer operators are often not sufficiently effective on their own, they are commonly combined with a smoothing procedure, leading to smoothing aggregation multigrid methods [3,15].

To the best of our knowledge, the use of segmentation algorithms within a multigrid or multilevel framework for imaging has not yet been explored. Moreover, they have not been used to define the restriction and prolongation operators

required in a multigrid approach. The possibility of defining grid transfer operators based on image segmentation provides a framework for exploring nonstandard properties of multigrid methods in imaging, beyond the image deblurring application considered in this paper. This approach can be interpreted as a data-dependent version of the multigrid framework, since the grid transfer operators are constructed from the observed image. This may lead to coarse representations of the original problem that are more meaningful than those obtained using standard restriction operators, regardless of the images in question.

In this work, we restrict our attention to the study of a cascadic multilevel approach developed in [12,13]. This approach differs from the classical multigrid framework in [4,7] in that the image is first restricted to a coarser level and, once an approximate solution has been computed using an iterative method, it is prolonged to the finer level, where further iterative refinement is performed by means of an iterative solver. Our proposal differs from that in [12,13] in the choice of the grid transfer operators, as we define both restriction and prolongation operators based on a segmentation obtained via a segmentation algorithm [1,9]. The main difficulty in the cascadic approach lies in the definition of suitable stopping criteria for the coarse-level problems. Here, we extend the analysis in [12,13] by proposing a suitable variant of the discrepancy principle.

This paper is organized as follows. Sections 2 and 3 introduce the two main components of our method: cascadic multigrid methods and segmentation algorithms. A detailed analysis of a variant of the discrepancy principle as a stopping criterion at the coarsest level is presented in Sect. 4. Section 5 is devoted to a set of numerical experiments demonstrating the robustness of the proposed two-level method, even in the case of aggressive coarsening.

2 Cascadic Multigrid

In [12], the authors proposed the use of a cascadic multilevel approach to solve image deblurring problems as (3). A multilevel method is characterized by a chain of nested subspaces $W_1 \subseteq W_2 \cdots \subseteq W_l = \mathbb{R}^n$. We will refer to the subspaces W_i as levels and we will say that W_i is coarser respect to W_{i+1} while W_{i+1} is finer respect to W_i. In order to move between the different levels, we need to define the restriction and prolongation operators. In a cascadic multiresolution approach, the restriction operator, usually denoted by R_i, is defined as $R_i : \mathbb{R}^n \to W_i$ while the prolongation operator is defined as $P_i : W_i \to W_{i+1}$.

The cascadic multilevel approach strategy aims to compute an approximate solution for the problem (3) at the coarsest level W_1 by means of an iterative solver. Then the prolongation operator P_1 is applied to this coarse solution to compute an initial guess for the level W_2. Once we are working in W_2, a correction of the initial approximation is computed and the iterative method is stopped, this new corrected solution is mapped to the finer level and corrected again. The procedure is presented in Algorithm 1.

Algorithm 1. Multilevel algorithm from [12]

Input: A, b^δ, δ, l the number of levels, an *IterativeMethod* and a stopping rule.
Output: An approximate solution $\tilde{x}$ in W_l

```
for i = 1, ..., l do
    compute A_i and b_i^δ
end for
x_0 = 0
for i = 1, ..., l do
    x_{i,0} = P_i x_{i-1}
    Δx_{i,k_i} = IterativeMethod(A_i, b_i^δ − A_i x_{i,0})
    correction step: x_i = x_{i,0} + Δx_{i,k_i}
end for
x̃ = x_l
```

The *IterativeMethod* is stopped at iteration k_i at the $i-$it level according to the stopping rule. It is applied to the error equation end hence $\Delta x_{i,k_i}$ is an approximation of the error at the $i-$th level.

We stress that this framework of the cascadic multigrid is independent of the choice of the restriction and prolongation operators. In this work, we explore the possibility of using restriction and prolongation operators defined starting from a segmentation of the observed image. Moreover, we consider only a 2-level cascadic approach. Therefore, we consider only two levels denoted as $W_c \subseteq W_f = R^n$. We will denote with $b_f^\delta, b_c^\delta, A_f$, and A_c, respectively, the observed signals and the operators at the coarse and fine level, respectively. Moreover, since we are dealing with only two levels, we will simply denote the restriction operator by R and the prolongation operator by P.

3 Segmentation and Coarsening

The output of any segmentation algorithm can viewed as a function that, given an input image $I \in \mathbb{R}^n$, provides a collection of elements $\mathcal{L} = \{\ell_1, \ldots, \ell_m\}$ and a function $\mathcal{P} : \mathcal{L} \times \mathbb{R}^n$ s.t.

$$\mathcal{P}(\ell_i, p) = \begin{cases} 1 \text{ if } p \in \ell_i \\ 0 \text{ otherwise,} \end{cases} \tag{6}$$

where $p \in \mathbb{R}^N$ are the pixels of the image I. The function (6) identifies if a pixel belongs to the element ℓ_i or not. For convenience, we denote the cardinality of ℓ_i by $\#\ell_i$ and the cardinality of the segmentation as $\#\mathcal{L} = m$.

We define the coarse level as $W_c = \mathbb{R}^{\#\mathcal{L}}$, which is the space of the elements of the segmentation of the observed image. Once we have a segmentation $\mathcal{L}$, we need to associate a value at each $\ell_i \in \mathcal{L}$. We select for this the mean of the values in the element. This choice resemble the standard multilevel choice of local

averaging on the pixels of the image.

More formally, the restriction operator $R : W_f \rightarrow W_c$ is defined as the average value of each pixel belonging to the same element of the segmentation. It is the product of the matrix $\Pi : \mathbb{R}^n \rightarrow \mathbb{R}^m$, which is the binary matrix that is 1 where $\mathcal{P}$ in Eq. (6) is, and the diagonal matrix Λ with diagonal entries $\Lambda_{i,i} = \text{Diag}(1/\#\ell_i)$.

Namely, we define the restriction operator as

$$R = \Lambda\Pi. \tag{7}$$

Fig. 1 show a visual representation of the application of the restriction operator R to an image.

In order to define a coarse representation A_c of the fine blurring operator $A_f = A$, we use a Galerkin approach such that

$$A_c = \Lambda\Pi A_f \Pi^T, \tag{8}$$

according to the choice $P = \Pi^T$.

The problem with this choice of P is that the prolonged image in W_f will result in a linear piecewise representation of the element of the segmentation. To avoid this problem, we define the prolongation operator as

$$P = G \star \Pi^T, \tag{9}$$

where $\star$ is the convolution operator and G is a Gaussian kernel of standard deviation 1. This has been performed to smooth the approximated solution on the fine grid, reducing the effect of the piecewise constant approximation that will result from the prolongation alone. With this new choice of the prolongation the coarser operator A_c in (8) is no longer associated to the Galerkin approach, which exhibits an increase in ill-conditioning at high frequencies. Anyway, the test performed with the Galerkin approach gives comparable results and hence we prefere to keep the coarser operator as A_c in (8), since it is cheaper to compute and better conditioned with respect to the true Galerkin operator.

Note that the prolongation P in Eq. (9) apply a smoothing step, by the Gaussian kernel, to the aggregation grid transfer operator. Indeed, a higher variance of the Gaussian kernel G can be interpreted a more steps of the smoothing of the aggregation prolongation [3]. Finally, when the segmentation algorithm provides a degenerate grid with a large variation in $\#\ell_i$, then the coarse operator A_c becomes more and more ill-conditioned leading to a decline in performance.

4 Stopping Criterion for Cascadic Multilevel

In order to find the approximate solution at every level, a stopping criterion is needed. This criterion must hold both for the fine and coarse levels. In [12,13]

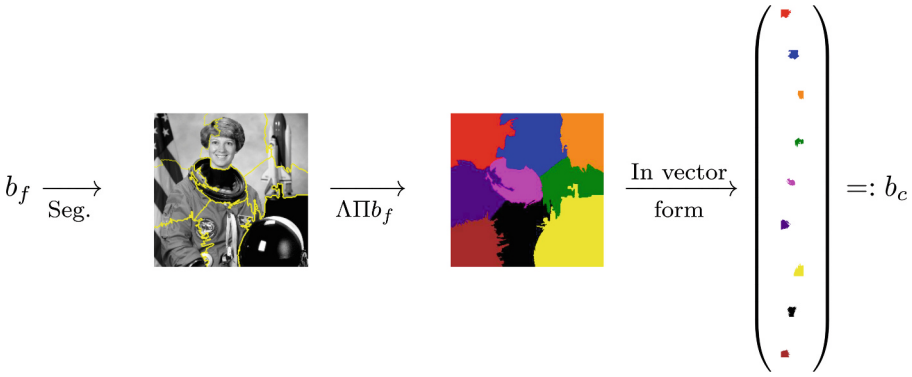


Fig. 1. Visual representation of the action of the restriction operator. Different colors represent different values.

the authors propose to use a variant of the Discrepancy Principle (DP) as a stopping rule at the coarse levels. We briefly recall the DP and the extension presented in the articles. Then, we further extend the stopping criterion to the coarsening based on a segmentation.

Given $v \in \mathbb{R}^n$, we define the weighted least squares norm as

$$\|v\| = \frac{1}{\sqrt{n}}\|v\|_2, \tag{10}$$

and we assume that we know an estimate of norm of the noise $\delta = \|b^\delta - b\|$, where b denotes the blurred but noise-free image. A vector x is said to satisfy the Discrepancy Principle (DP) if $\|Ax - b^\delta\| < \gamma\delta$, for a fixed $\gamma > 1$. Given an iterative method, the stopping rule based on the DP terminates the iterative process at the first step k such that

$$\|Ax_k - b^\delta\| < \gamma\delta. \tag{11}$$

Starting from this stopping criterion, the authors propose to extend it to the cascadic multilevel setting.

In Algorithm 1, the *IterativeMethod* is applied to the linear system $A\Delta x_i = b_i^\delta - A_i x_{i,0}$ and stopped by DP. In order to stop the iteration at the i-th level, let c_i be such that $\|b_i^\delta - b_i\| \leq c_i\delta$, we terminate the iteration of the *IterativeMethod* at level i when

$$\|b_i^\delta - A_i x_{i,0} - A_i \Delta x_{i,k_i}\| \leq c_i\delta, \quad i = 1, \ldots, l. \tag{12}$$

where $c_l = \gamma$ and k_i is the $k-$th iteration at the $i-$th level. In the case where the restriction operator coincides with the weighted average of four adjacent pixels, the authors [12] proved that the relation $c_i = \frac{1}{2}c_{i+1}$ holds.

4.1 Estimation of c_1 in the Segmentation Case

Since we are considering a 2-level method, we need to determine a constant c_1 that must hold at the coarse level. To compute this constant, first consider the coarse residual $e_c = b_c^\delta - b_c$. We define the average variance norm of e_c given a segmentation $\mathcal{L}$ as

$$\|e_c\|_V^2 = \frac{1}{\#\mathcal{L}} \sum_{\ell_i \in \mathcal{L}} \mathrm{Var}(e[\ell_i]), \tag{13}$$

where $e[\ell_i] = \frac{1}{\#\ell_i} \sum_{p \in \ell_i} e^p$ is the mean with respect to the element ℓ_i and e^p is the $p-$th entry of e. Therefore, by the properties of the variance, we have

$$\|e_c\|_V^2 = \frac{1}{\#\mathcal{L}} \sum_{\ell_i \in \mathcal{L}} \frac{1}{\#\ell_i^2} \sum_{p \in \ell_i} \mathrm{Var}(e^p). \tag{14}$$

By observing that $\sum_{\ell_i \in \mathcal{P}} \sum_{p \in \ell_p} \mathrm{Var}(e^p) = \sum_{i=1}^n \mathrm{Var}(e^i)$, we obtain

$$\frac{n}{\#\mathcal{L} \cdot \max(\ell_i)^2} \|e_f\|_V^2 \quad \leq \|e_c\|_V^2 \leq \quad \frac{n}{\#\mathcal{L} \cdot \min(\ell_i)^2} \|e_f\|_V^2. \tag{15}$$

Using the approximation $\|e_f\|_V \approx \delta$, we define

$$c_{max} = \gamma \sqrt{\frac{n}{\#\mathcal{L} \max(\ell_i)^2}}, \quad c_{min} = \gamma \sqrt{\frac{n}{\#\mathcal{L} \min(\ell_i)^2}}, \quad c_{avg} = \frac{c_{min} + c_{max}}{2}. \tag{16}$$

We observe that c_{min} and c_{max} in (16) depend on the segmentation; in particular, they depend only on the minimum and maximum number of pixels present in an element of the segmentation. The choice $c_1 = c_{min}$ is consistent with the stopping criterion (12). In practice, $c_1 = c_{avg}$ usually provides a good choice in the numerical experiments.

Example 1. If we consider the case where the restriction operator is the average of four adjacent pixels in the image, we have that $\#\mathcal{L} = n/4$ and $\min(\#\ell_i) = \max(\#\ell_i) = 4$, thus Eq (15) leads to

$$\frac{1}{4}\|e_f\|_V^2 \leq \|e_c\|_V^2 \leq \frac{1}{4}\|e_f\|_V^2. \tag{17}$$

Thus, the same conclusion obtained in [12] that $c_1 = \frac{1}{2} c_f$ holds.

5 Numerical Experiments

In this section, we present three numerical experiments. For all the numerical experiments, we consider the 2-level method proposed. Firstly, we obtain the initial segmentation computed on the corrupted observed image by means of two well-known segmentation algorithms: SLIC [1] and Felzenszwalb's algorithm [9]. For both algorithms, we use the versions implemented in the scikit-image library [17]. The Landweber method is used as an iterative solver in Algorithm 1. We consider Algorithm 1 where, after the segmentation is obtained, we project the

image on the coarse level, then we obtain a reconstruction at the coarse level by means of Landweber, we prolong it to the fine level, and correct it.

In each test, we consider a test image that is blurred using two different kernels, then additive white Gaussian noise is added to the image. We add the noise ξ such that b^δ is computed as

$$b^\delta = b + \delta\xi \tag{18}$$

and we denote by δ the noise level. The parameter γ for the Discrepancy Principle (11) is set to 1.01 while, for the Landweber steplenght α_k we use a constant choice of $\alpha_k = \frac{1}{\|A\|_2^2}$ at the fine level and $\alpha_k = \frac{1}{\|A_c\|_2^2}$ and at the coarse level. In the last numerical experiment, we use CLGS instead of Landweber as iterative solver.

5.1 First Numerical Experiment: Gaussian Blur and 0.5% of Noise Level

We consider the 256×256 test image in Fig. 2. The true image is corrupted by the convolution with a Gaussian kernel with a standard deviation of 5 and additive white Gaussian noise of level 0.5%. The cropped PSF, i.e., the convolution operator applied to a point, is reported in Fig. 2. The segmentation is obtained by applying the SLIC algorithm directly on the observed corrupted image. The resulting segmentation consists of 53 elements.

In Fig. 3, we report the approximations obtained via the proposed 2-level method and the one obtained with the standard Landweber method. For comparison, we report the value of the PSNR and the SSIM for both images in the caption.

In Fig. 4, the evolution of the PSNR against the number of iterations is shown for both the two-level method and the Landweber method. We observe that the proposed method consistently produces approximations with higher PSNR values. Moreover, the standard Landweber method requires approximately twice as many iterations to achieve comparable PSNR values. The number of iterations performed at the coarse level is 5, which is negligible compared to the number of fine-level iterations.

5.2 Second Numerical Experiment: Box Blur and 2% of Noise Level

As a second test, we consider the 256×256 test image in Fig. 5, the image is corrupted by a box blur and a noise level of 2% is added. The corrupted image and the PSF are reported in Fig. 5. In this test, we consider Felzenszwalb's algorithm for the segmentation. This algorithm gives rise to segmentations that are less regular with respect to the segmentations obtained using the SLIC algorithm, as we can observe in the segmentation presented in Fig. 5.

The resulting segmentation consists of 126 elements. In Fig. 6, the approximated solutions obtained by Landweber and the proposed method are presented.

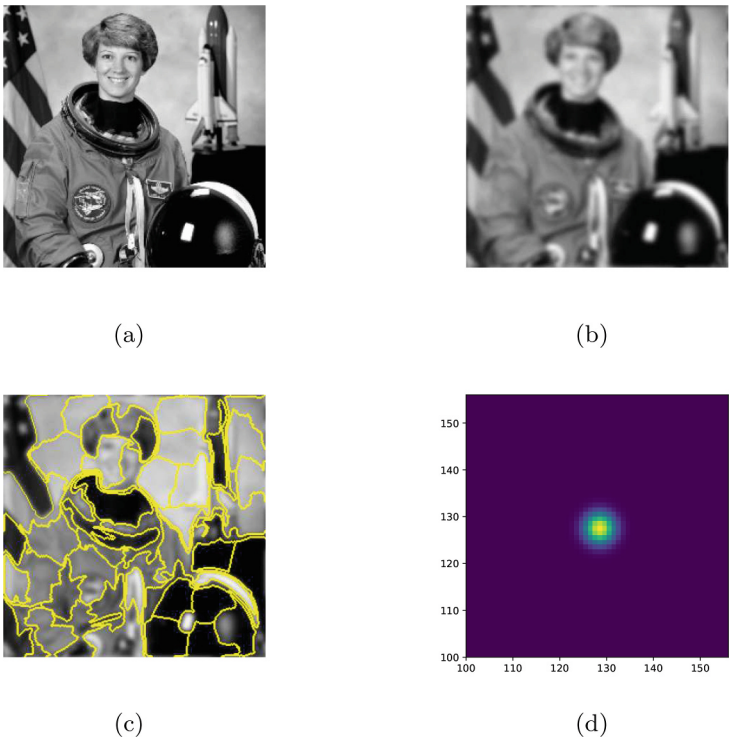


Fig. 2. (a) Original test image, (b) corrupted observed image, (c) segmentation used, (d) PSF used in the test.

Figure 7 shows the evolution of the performance metric against the number of iterations. The iterations performed at the coarse level are 2.

According to the results of the first experiment, we observe that the cascadic 2-level method improves the reconstruction and, in a number of fine iterations that are comparable to standard Landweber, returns an approximation with higher value of PSNR. If we consider the evolution of the iterations, we observe that the 2-level method obtains reconstructions that are with a better PSNR value.

5.3 Third Numerical Experiment: Test with Different Segmentation Sizes

As an additional test, we investigate the effect of the starting segmentation on the evolution of the performance metrics. We compute three initial segmentations using the SLIC algorithm that are of different sizes: 62, 248, and 521. In this test, we are considering a noise level of 1% and a Gaussian blur of standard deviation 3. In Fig. 8 we present the three initial segmentations proposed, the PSF, and the evolution of the PSNR for the 2-level cascadic algorithm as well

(a)

(b)

Fig. 3. (a) Landweber method PSNR: 23.12 SSIM: 0.792, (b) 2-level cascadic PSNR: 23.32 SSIM: 0.795

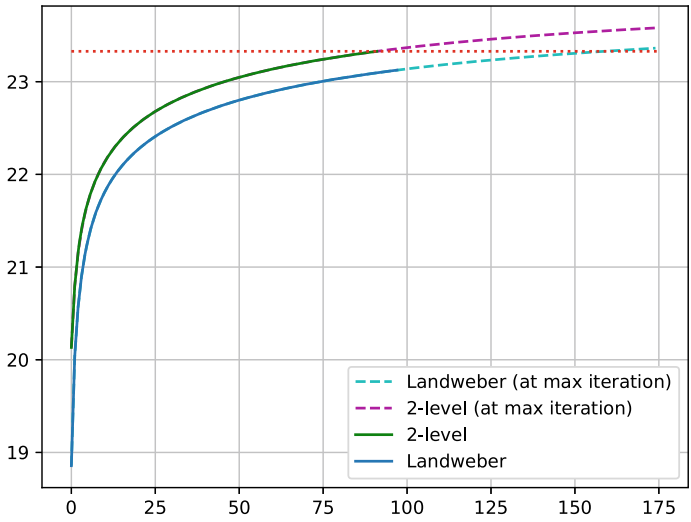


Fig. 4. PSNR value against the iteration number. The red dotted line represents the PSNR value reached for the 2-level method. (Color figure online)

as for standard Landweber. We observe that using a larger number of segmentation elements leads to higher PSNR values. However, the improvement over segmentations with fewer elements is marginal.

5.4 Fourth Numerical Experiment: Test Using CGLS Instead of Landweber

To prove the flexibility of our 2-level approach, we replace the Landweber iteration with the CGLS method [2] in Algorithm 1. We perform the same experiment as in Subsect. 5.1 cf. Fig. 2. As in the previous test, a noise level of 0.5% is added. The final reconstructions obtained with both CGLS and the 2-level

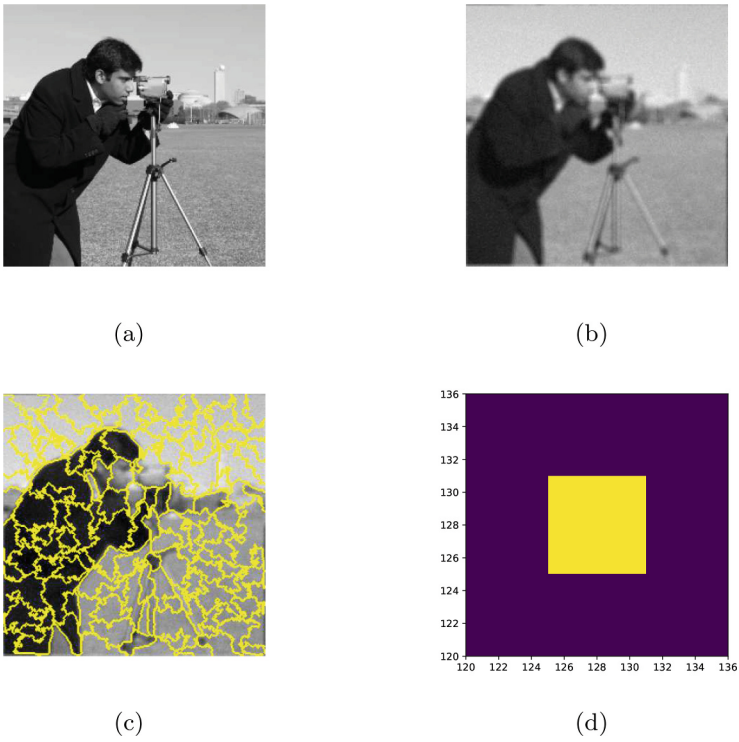


Fig. 5. (a) Original test image, (b) corrupted observed image, (c) segmentation used, (d) PSF used in the test.

method with CGLS as standard iterative solver are reported in Fig. 9. The corresponding performance metrics are reported in the caption. The number of iterations performed at the coarse level is 4.

In Fig. 10, we report the evolution of the performance metrics against the iteration number. As expected, the CGLS method converges faster than the Landweber method. Moreover, in accordance with the previous tests, the proposed 2-level method produces reconstructions with higher PSNR values in fewer iterations with respect to the standard iterative method.

(a)

(b)

Fig. 6. (a) Landweber method PSNR: 23.87 SSIM: 0.652, (b) 2-level cascadic PSNR: 24.02 SSIM: 0.657

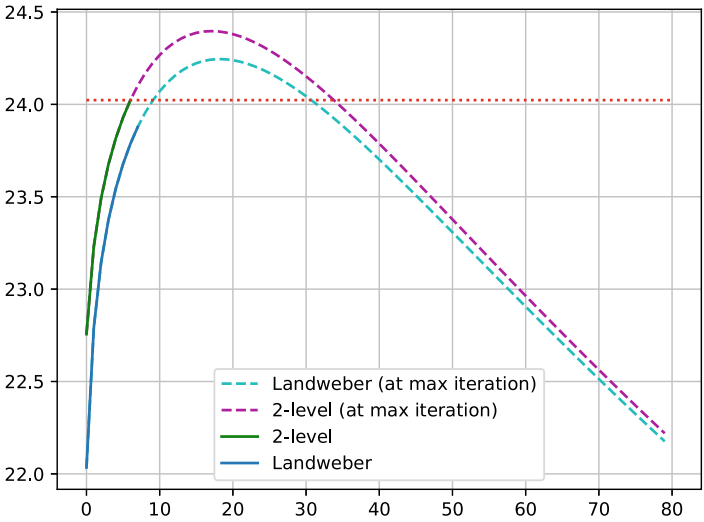
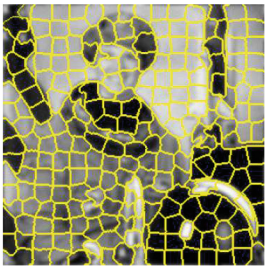
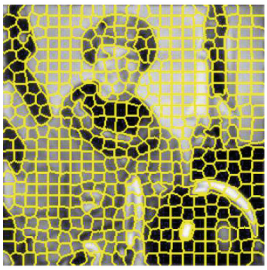


Fig. 7. PSNR value against the iteration number. The red dotted line represents the PSNR value reached for the 2-level method. (Color figure online)

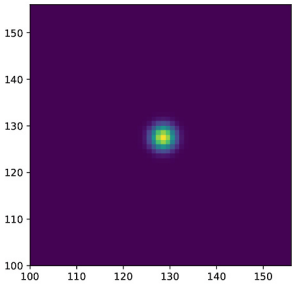
(a)



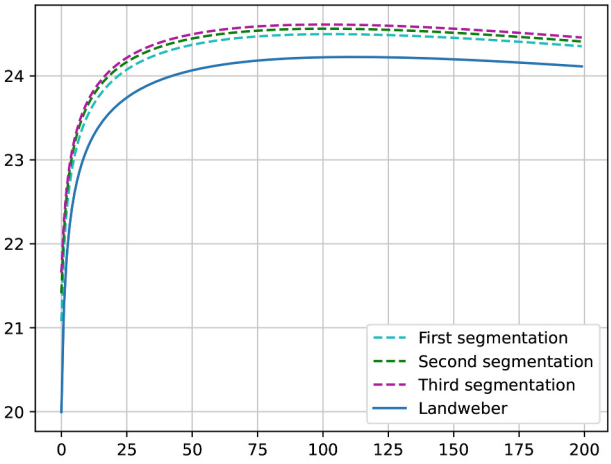
(b)



(c)



(d)



(e)

Fig. 8. (a) First segmentation of 62 elements, (b) second segmentation of 248 elements, (c) third segmentation of 521 elements, (d) evolution of the PSNR for the three segmentations and the standard Landweber against the iteration number.

(a) (b)

Fig. 9. (a) CGLS solution PSNR: 23.18, SSIM: 0.782, (b) 2-level solution PSNR: 23.39, SSIM: 0.788

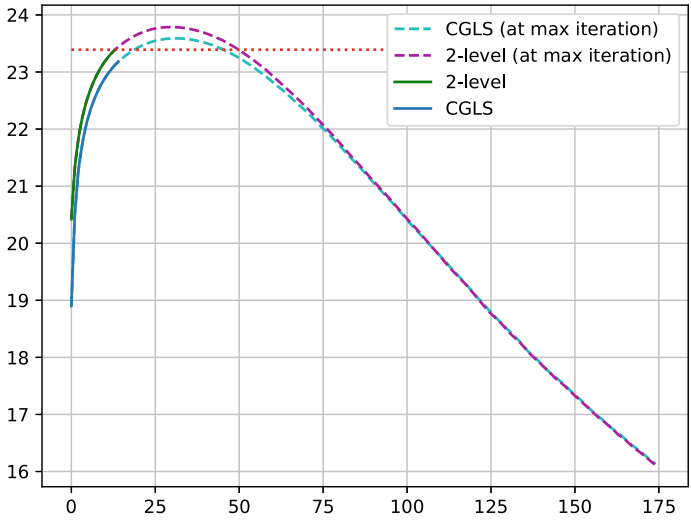


Fig. 10. Evolution of the PSNR against the number of iterations. The red dotted line represents the PSNR value reached for the 2-level method. (Color figure online)

6 Conclusions

In this work, we investigated the possibility of using nonstandard, segmentation-aware grid transfer operators in a cascadic multigrid setting. We then derived a stopping criterion for the coarse level. The numerical experiments with different blurring operators and noise levels confirm the robustness of the proposed two-level method, even under aggressive coarsening projections. The initial computation of the projection operator necessarily introduces computational overhead. This overhead depends on both the time complexity of the initial segmentation and the cost of the projection at the coarse level. The time complexity for the initial segmentation is $O(N)$ for SLIC [1] and $O(N \log N)$ for Felzenszwalb's

algorithm [9]. In addition, the cost of projecting the problem at the coarse level must be taken into account. The use of these operators in a non-cascadic multilevel setting remains to be explored, as does their extension to more than two levels. Moreover, the adoption of other iterative solvers, such as the iterative soft-thresholding algorithm and its accelerated versions [5,11], should be considered. The use of an iterative soft-thresholding algorithm, which aims to promote sparsity with respect to a given basis (e.g., wavelets) at the coarse level, inherently requires the definition of both the projection onto the chosen basis and the optimization problem to be solved at that level.

Acknowledgments. The authors thank Professor Rolf Krause for the valuable discussions that contributed to this work, in particular for his insights into multigrid algorithms. The authors are partially supported by the Italian Ministry of the University and Research, through the PRIN 2022 project "Inverse Problems in the Imaging Sciences (IPIS)" (No. 2022ANC8HL), and GNCS–INdAM.

Disclosure of Interests. The authors have no competing interests.

References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Pattern Anal. Mach. Intell. **34**(11), 2274–2282 (2012)
2. Björck, Å.: Numerical Methods for Least Squares Problems, 2nd edn. Society for Industrial and Applied Mathematics, Philadelphia, PA (2024)
3. Bolten, M., Donatelli, M., Huckle, T., et al.: Analysis of smoothed aggregation multigrid methods based on Toeplitz matrices. Electron. Trans. Numer. Anal. **44**, 25–52 (2015)
4. Buccini, A., Donatelli, M.: A multigrid frame based method for image deblurring. Electron. Trans. Numer. Anal. **53**, 283–312 (2020)
5. Daubechies, I., Defrise, M., De Mol, C.: An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Commun. Pure Appl. Math. A J. Issued by the Courant Institute Math. Sci. **57**(11), 1413–1457 (2004)
6. Donatelli, M.: A multigrid for image deblurring with Tikhonov regularization. Numer. Linear Algebra Appl. **12**(8), 715–729 (2005)
7. Donatelli, M., Serra-Capizzano, S.: On the regularizing power of multigrid-type algorithms. SIAM J. Sci. Comput. **27**(6), 2053–2076 (2006)
8. Engl, H.W., Hanke, M., Neubauer, A.: Regularization of inverse problems, vol. 375. Springer Science & Business Media (1996)
9. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient graph-based image segmentation. Int. J. Comput. Vis. **59**(2), 167–181 (2004)
10. Hansen, P.C., Nagy, J.G., O'Leary, D.P.: Deblurring images: matrices, spectra, and filtering. SIAM (2006)
11. Huang, J., Donatelli, M., Chan, R.H., et al.: Nonstationary iterated thresholding algorithms for image deblurring. Inverse Probl. Imaging **7**(3), 717–736 (2013)
12. Morigi, S., Reichel, L., Sgallari, F.: Noise-reducing cascadic multilevel methods for linear discrete ill-posed problems. Numer. Algorithms **53**(1), 1–22 (2010)
13. Morigi, S., Reichel, L., Sgallari, F., Shyshkov, A.: Cascadic multiresolution methods for image deblurring. SIAM J. Imaging Sci. **1**(1), 51–74 (2008)

14. Notay, Y.: An aggregation-based algebraic multigrid method. Electron. Trans. Numer. Anal. **37**(6), 123–146 (2010)
15. Vaněk, P., Brezina, M., Mandel, J.: Convergence of algebraic multigrid based on smoothed aggregation. Numer. Math. **88**(3), 559–579 (2001)
16. Vaněk, P., Mandel, J., Brezina, M.: Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems. Computing **56**(3), 179–196 (1996)
17. van der Walt, S., et al.: scikit-image contributors: scikit-image: image processing in Python. PeerJ **2**, e453 (2014)

Neural Network and Sampling Kantorovich Operators vs Traditional Denoising Techniques

Ana Maria Acu[1(✉)], Cristian Daniel Alecsa[1], Arianna Travaglini[2], and Gianluca Vinti[1,3]

[1] Department of Mathematics and Informatics, Lucian Blaga University of Sibiu, 5-7, Str. Dr. I. Raţiu, 550012 Sibiu, Romania
anamaria.acu@ulbsibiu.ro, gianluca.vinti@unipg.it

[2] Department of Physics and Geology, University of Perugia, Via A. Pascoli, 06123 Perugia, Italy
arianna.travaglini@collaboratori.unipg.it

[3] Department of Mathematics and Computer Science, University of Perugia, 1, Via Vanvitelli, 06123 Perugia, Italy

Abstract. This paper addresses the problem of image reconstruction in the presence of different types of noise, focusing on the analysis and comparison of several filtering techniques. In particular, Neural Network (NN) operators and sampling Kantorovich (SK) operators are investigated alongside a variety of classical and advanced filtering methods, including linear, nonlinear, and state-of-the-art approaches commonly adopted in the literature. The performance of the considered methods is evaluated through widely used quantitative metrics. The experimental analysis highlights the ability of NN and SK operators to provide accurate and stable reconstructions, ensuring the effectiveness and versatility of the proposed operators across different noise conditions.

Keywords: Neural network operators · sampling Kantorovich operators · Image filtering · Image smoothing

1 Introduction

In recent years, the study of advanced approximation techniques for image processing has attracted significant attention, particularly in relation to operator-based methods and their applications to image reconstruction, rescaling, and denoising. These issues are especially critical in domains where accurate visual information is essential, such as medical imaging, remote sensing, and scientific visualization, where reliable data is crucial for analysis, interpretation, and decision-making.

Over the years, a variety of filtering techniques have been developed to reduce noise and improve the overall quality of images. Each filter is generally designed to address specific types of noise—for instance, Gaussian filters are commonly

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 124–141, 2027.
https://doi.org/10.1007/978-3-032-30530-5_8

applied to mitigate Gaussian noise. However, the limited adaptability of these classical filters to different noise models represents a significant drawback, motivating the development of more flexible and general approaches.

Among the most promising approaches are neural network (NN) operators, which have been extensively investigated due to their strong connections with the theory of artificial neural networks (see [7–10,15,20]). The NN operators activated by a sigmoidal function σ and acting on $f : I \to \mathbb{R}$ are defined as

$$F_n(f, \mathbf{x}) := \frac{\sum\limits_{k \in \mathcal{A}_n} f\left(\frac{\mathbf{k}}{n}\right) \Psi_\sigma(n\mathbf{x} - \mathbf{k})}{\sum\limits_{k \in \mathcal{A}_n} \Psi_\sigma(n\mathbf{x} - \mathbf{k})},$$

for every $\mathbf{x} \in I$ and $\mathbf{k} = (k_1, k_2, \ldots, k_d)$, where $I := \prod\limits_{i=1}^{d} [a_i, b_i]$ is a d-dimensional interval belonging to $\mathbb{R}^d$, $\mathcal{A}_n := \{\mathbf{k} \in \mathbb{Z}^d : \lceil na_i \rceil \leq k_i \leq \lfloor nb_i \rfloor \text{ for } i = 1, \ldots, d\}$, and where the symbols $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denote the floor and ceiling functions, respectively (see, e.g., [14]).

Alongside NN-based approaches, sampling Kantorovich (SK) operators represent another powerful tool for image reconstruction and approximation. Originally introduced in [4] in the univariate setting and later extended to the multidimensional case in [13], these operators are defined as

$$(SK_w f)(\mathbf{x}) := \sum_{\mathbf{k} \in \mathbb{Z}^d} \chi(w\mathbf{x} - \mathbf{k}) \left[w^d \int\limits_{R_{\mathbf{k}}^w} f(\mathbf{u})\, d\mathbf{u} \right], \quad \mathbf{x} \in \mathbb{R}^d,$$

where χ is the kernel of the operator and $f : \mathbb{R}^d \to \mathbb{R}$ is a locally integrable function such that the series converges for every $\mathbf{x} \in \mathbb{R}^d$. SK operators provide a powerful generalization of the classical sampling framework, extending it to functions that may be discontinuous, and have been studied in various functional settings, including Orlicz spaces. They have proven particularly effective in handling real-world imaging problems (see, e.g., [11,22,25,27,28]). Indeed, both SK and NN operators are used as smoothing filters due to their low-pass filtering behavior given by their convolutive nature (see [1]). To evaluate their effectiveness and versatility in reconstructing images corrupted by various types of noise, we perform a comparative analysis considering three different and commonly encountered types of noise—Gaussian, Poisson, and Speckle—to ensure that the comparison reflects the methods' adaptability across different noise conditions (see [19,31]). The reconstructed images obtained through SK and NN operators are then compared with those produced by classical image filtering techniques, such as Gaussian, Wiener, BM3D, Median, Bilateral, Non-Local Means (NLM), a combination of Variance Stabilizing Transform (VST) and Non-Local Means, Lee, Frost, Log-Wiener and Wiener-Anscombe filters (see, e.g., [3,6,16–18,21,23,24,26,30]).

To quantitatively assess the quality of the reconstructed images, several widely used similarity measures are employed, including the Root Mean Square

Error (RMSE), the Structural Similarity Index (SSIM), the Feature Similarity Index (FSIM), and the Feature-based Loss Index for Perceptual quality (ꟻLIP) (see [2,5,29,32]). Unlike classical pixel-wise error metrics such as RMSE, SSIM captures structural information by accounting for luminance, contrast, and local correlations. These metrics provide complementary measures of image fidelity, perceptual quality, and structural consistency, allowing for a comprehensive evaluation of the denoising performance of each method.

Overall, this work underline the effectiveness of operator-based approximation methods, such as NN and SK operators, in modern image processing. By integrating theoretical foundations from approximation theory with practical algorithmic design and evaluation through established metrics, these approaches provide robust and flexible tools for tackling complex tasks such as image denoising and reconstruction across different noise models.

2 Algorithmic Framework

This section presents the mathematical and algorithmic framework for image reconstruction and resizing based on multivariate neural network (NN) and sampling Kantorovich (SK) operators. Both approaches are formulated in a continuous setting and can be naturally applied to grayscale images interpreted as functions over multidimensional domains.

The multivariate NN operator considered, activated by a sigmoidal function σ, has been studied, e.g., in [1,14] and it is defined for $n \in \mathbb{N}^+$, as

$$F_n(f,\mathbf{x}) := \frac{\sum\limits_{k\in\mathcal{A}_n} f\left(\frac{\mathbf{k}}{n}\right)\Psi_\sigma(n\mathbf{x}-\mathbf{k})}{\sum\limits_{k\in\mathcal{A}_n}\Psi_\sigma(n\mathbf{x}-\mathbf{k})}, \quad \frac{\mathbf{k}}{n} := \left(\frac{k_1}{n},\ldots,\frac{k_d}{n}\right), \quad \text{(NN_Operator)}$$

where $f : I \to \mathbb{R}$, with $I := \prod\limits_{i=1}^{d}[a_i,b_i]$ is a d-dimensional interval belonging to $\mathbb{R}^d$, and $\mathcal{A}_n := \{\mathbf{k}\in\mathbb{Z}^d : \lceil na_i\rceil \le k_i \le \lfloor nb_i\rfloor \text{ for } i=1,\ldots,d\}$. In addition, the sigmoidal function $\sigma : \mathbb{R}\to\mathbb{R}$ satisfies $\lim\limits_{x\to-\infty}\sigma(x)=0$ and $\lim\limits_{x\to+\infty}\sigma(x)=1$, along with the following assumptions:

(Σ_1) the function $x \to \sigma(x) - \frac{1}{2}$ is odd.
(Σ_2) $\sigma \in C^2(\mathbb{R})$ is concave for $x \ge 0$.
(Σ_3) $\sigma(x) = \mathcal{O}\left(\frac{1}{|x|^{1+\alpha}}\right)$ as $x \to -\infty$, for some $\alpha > 0$.

The density function associated to σ is defined, for $x \in \mathbb{R}$, as

$$\phi_\sigma(x) := \frac{1}{2}\left[\sigma(x+1) - \sigma(x-1)\right]$$

and the multivariate kernel function Ψ_σ is constructed by means of the tensor product of the density functions, namely

$$\Psi_\sigma(\mathbf{x}) := \prod_{i=1}^{d} \phi_\sigma(x_i).$$

Within this framework, the NN operator reconstructs the function f by means of normalized weighted combinations of sampled values. Each reconstructed point is obtained as a weighted average of sampled values, where the weights are determined by the kernel Ψ_σ, thus preserving local image features while ensuring numerical stability. Note that, due to a mere typo, in [1] the rectified linear unit function (ReLU) is denoted using the symbol of a sigmoidal function, although it is not one.

An alternative reconstruction strategy is provided by SK operators introduced in [4], originally to deal with one-dimensional discontinuous signals and recently studied in a modified version to better address the problem of signals affected by noise (see [12]). The multivariate generalization of these operators was proposed in [13], which allows the application of the framework of SK operators to image processing. In this setting, the family of sampling Kantorovich operators $(SK_w)_{w>0}$ is given by

$$(SK_w f)(\mathbf{x}) := \sum_{\mathbf{k}\in\mathbb{Z}^d} \chi(w\mathbf{x}-\mathbf{k}) \left[w^d \int_{R_{\mathbf{k}}^w} f(\mathbf{u})\, d\mathbf{u} \right], \quad \mathbf{x}\in\mathbb{R}^d,\ w>0, \tag{SK_operator}$$

where $f : \mathbb{R}^d \to \mathbb{R}$ is a locally integrable, representing the signal of the image, such that the series converges for all $\mathbf{x} \in \mathbb{R}^d$. Here, the sets

$$R_{\mathbf{k}}^w := \prod_{i=1}^{d} \left[\frac{k_i}{w}, \frac{k_i+1}{w} \right]$$

denote the regions over which the local averages of f are computed. In addition, the kernel function $\chi : \mathbb{R}^d \to \mathbb{R}$ is required to satisfy the following conditions:

- (χ_1) The kernel χ belongs to $L^1(\mathbb{R}^d)$ and it is bounded in a neighborhood of the origin of $\mathbb{R}^d$.
- (χ_2) The kernel forms a partition of unity, i.e., $\sum_{\mathbf{k}\in\mathbb{Z}^d} \chi(\mathbf{x}-\mathbf{k}) = 1$ for all $\mathbf{x}\in\mathbb{R}^d$.
- (χ_3) There exists $\beta > 0$ such that the discrete absolute moment of order $\beta > 0$ is finite, namely $m_\beta(\chi) := \sup_{\mathbf{u}\in\mathbb{R}^d} \sum_{\mathbf{k}\in\mathbb{Z}^d} |\chi(\mathbf{u}-\mathbf{k})| \cdot \|\mathbf{u}-\mathbf{k}\|_2^\beta < +\infty$.

A fundamental difference between the NN operator and the SK operator mainly relies on the fact that, in the latter mentioned, the pointwise samples are replaced by the integral mean over each region, which smooths local variations and reduces time-jitter errors. Therefore, the SK algorithm introduces inherent regularization, yielding robust reconstructions, especially in the presence of noise. Therefore, the SK operator offers a stable and regularized reconstruction framework with respect to the images corrupted with different types of input noise.

For implementation purposes, a grayscale image can be interpreted as a matrix $A = (a_{ij})_{i,j=1,\ldots,m}$, by means of which it is possible to construct the so called image function I as:

$$I(x,y) := \sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij}\mathbf{1}_{ij}(x,y)$$

for every $(x,y) \in \mathbb{R}^2$, where the indicator function $\mathbf{1}_{ij}$ is defined as

$$\mathbf{1}_{ij}(x,y) = \begin{cases} 1, & \text{if } (x,y) \in (i-1,i] \times (j-1,j], \\ 0, & \text{otherwise.} \end{cases}$$

Therefore, the function I (serving as f) can be approximated using two-dimensional NN operators as well as SK operators. This can be achieved by appropriately choosing a sigmoidal function in the NN setting and a kernel function in the SK setting, as discussed earlier. The proposed methods have been implemented in MATLAB©, providing a computational framework for image reconstruction and resizing.

The implementation of the NN operator described in Algorithm 1 begins by taking a grayscale image and preparing it for upscaling or reconstruction, based on the chosen refinement factor n and scale factor R. The input image is first replicated using the Kronecker tensor product, considering the parameter n, in order to create a fine-grid intermediate representation. A sampling grid is then generated, and a kernel function is evaluated to determine the contribution of the neighboring pixel. The algorithm identifies the effective kernel support by including only the non-zero terms. Finally, the creation of image and kernel slices are created while respecting image boundaries, and their weighted contributions are combined to produce the upscaled or reconstructed output image. In contrast, the SK-based implementation, in which the reconstruction process

Algorithm 1 `NN_Operator`

Objective: Reconstructing or resizing the original image I_{in} by means of the multivariate *NN operator* based on kernels constructed using sigmoidal functions.

Inputs: Original image I_{in} of size $H_{in} \times W_{in}$, parameter $n \in \mathbb{N}^+$, scaling factor $R \geq 1$.

- ▷ Choice and definition of the kernel Ψ_σ.
- ▷ Size of the output image: $H_{out} \times W_{out}$, where $H_{out} = R \cdot H_{in}$ and $W_{out} = R \cdot W_{in}$.
- ▷ Computation of the matrices of the sample values.
- ▷ Definition of the vectors containing the arguments of Ψ_σ.
- ▷ Definition of the reconstruction grid.

Iteration: Summation over $\mathbf{k}$ of all non-zero terms of the form $\Psi_\sigma(n\mathbf{x}-\mathbf{k})I_{in}\left(\frac{k}{n}\right)$, normalized by $\Psi_\sigma(n\mathbf{x}-\mathbf{k})$, for each point x belonging to the reconstruction grid, according to the formula (NN_Operator).

Output: Reconstructed image I_{out} of size $H_{out} \times W_{out}$.

Algorithm 2 `SK_Operator`

Objective: Reconstructing or resizing the original image I_{in} by means of the multivariate *SK operator* constructed using kernel functions.

Inputs: Original image I_{in} of size $H_{in} \times W_{in}$, parameter $w \in \mathbb{N}^+$, scaling factor $R \geq 1$.

- ▷ Choice and definition of the kernel χ.
- ▷ Size of the output image: $H_{out} \times W_{out}$, where $H_{out} = R \cdot H_{in}$ and $W_{out} = R \cdot W_{in}$.
- ▷ Computation of the matrices of the sample values.
- ▷ Definition of the vectors containing the arguments of χ.
- ▷ Definition of the reconstruction grid.

Iteration: Summation over $\mathbf{k}$ of all non-zero terms of the form $\chi(w\mathbf{x} - \mathbf{k}) \left[w^d \int_{R^w_{\mathbf{k}}} I_{in}(\mathbf{u})\, d\mathbf{u} \right]$, for each point $\mathbf{x}$ belonging to the reconstruction grid, according to the formula (SK_operator).

Output: Reconstructed image I_{out} of size $H_{out} \times W_{out}$.

is described in Algorithm 2, computes local averages of the image over subregions and combines them using kernel-based weights. This procedure naturally incorporates smoothing and leads to improving stability in noisy environments.

3 Experimental Data

The test images adopted in this paper are: "autumn.png" (345 × 206), "cameraman.png" (256 × 256), "lena.png" (200 × 200), "moon.png" (358 × 537), and "onion.png" (198 × 135). These images were selected to exhibit different characteristics such as size, brightness, texture, and structural complexity, allowing for a comprehensive evaluation of the denoising algorithms under varied conditions.

To study the filtering process and assess the performance of NN and SK operators, the original images were corrupted with three types of noise commonly encountered in real-world applications: Gaussian, Poisson, and Speckle noise. The noise corruption was implemented using MATLAB® with the following parameters: Gaussian noise with mean 0 and variance 0.05, Speckle noise with variance 0.04, and Poisson noise with standard parameters. Figure 1 presents both the original images and their corrupted versions for each noise type.

Fig. 1. Original test images (first column) and their corrupted versions with Gaussian (second column), Poisson (third column), and Speckle noise (fourth column).

4 NN and SK Optimization

After generating the noisy versions of the reference images, the next step was to optimize the filtering operators to achieve the best possible restoration of the original images. Since both SK and NN operators are parameter-dependent, their performance is strongly determined by the specific combination of parameters used. To enable comparisons, the scaling factor R has been set to 1 for both NN and SK algorithms.

For the NN operators, we tested all activation functions and network parameters to identify configurations that minimize the reconstruction error while preserving image structures. The density functions considered (see, e.g., [1]), indexed by the parameter k, are those ones constructed by using the following sigmoidal functions: hyperbolic tangent function ($k = 1$), logistic function ($k = 2$), ramp function ($k = 3$) and B-spline-based functions of orders 2–4 ($k = 4, 5, 6$). The parameter n controls the density of sampling points in the operator.

For the SK operators, the bandwidth parameter w determines the degree of smoothing applied to the image. The kernel used in all experiments is the

Jackson-type kernel, which proved to be well-suited for applications (see, e.g., [11,25,27]).

To identify the optimal parameters, we evaluated multiple quality indices (SSIM, RMSE, FSIM, and ꟻLIP) for each filtered image by comparing it with the original version. The images were normalized to the dynamic range $[0, 1]$, so that the results of all indices belong to the interval $[0, 1]$ too. To ensure consistency when computing an overall average, we considered $1-$RMSE and $1-$ꟻLIP so that higher values uniformly indicate better performance across all metrics. We then calculated the mean of these four normalized indices, referred to as the *index mean*. The optimal parameters for each image and noise type are those that maximize this index mean.

Table 1 summarizes the optimal parameter values obtained for each test image and noise type combination.

Table 1. Optimal parameters for SK and NN operators across different images and noise types.

Image	Noise Type	SK: w	NN: n	NN: k
Cameraman	Gaussian	4	1	2
	Poisson	16	3	2
	Speckle	9	2	2
Lena	Gaussian	5	1	2
	Poisson	17	4	2
	Speckle	10	2	2
Moon	Gaussian	4	1	2
	Poisson	11	2	2
	Speckle	6	1	2

From the optimization results, several patterns emerge. For NN operators, the logistic activation function ($k = 2$) consistently yields the best performance across all images and noise types, with the optimal value of n ranging between 1 and 4 depending on the specific noise characteristics. Lower values of n (typically $n = 1$) are optimal for Gaussian noise, while Poisson noise generally requires slightly higher values ($n = 2$ to $n = 4$). Speckle noise is optimally filtered for $n = 1$ or $n = 2$.

For SK operators, the optimal bandwidth parameter w exhibits greater variability depending on the noise type. Gaussian noise is best handled with relatively low values ($w = 4$ or $w = 5$), which provide sufficient smoothing without excessive blurring. Poisson noise, characterized by signal-dependent variance, requires higher values of w (ranging from $w = 11$ to $w = 17$) to achieve effective denoising. Speckle noise shows intermediate optimal values ($w = 6$, $w = 9$ or $w = 10$). These results are consistent with the theoretical requirements: values of w that are too high cause the reconstruction to converge toward the noisy

image, while values that are too low produce excessively blurred results with insufficient noise reduction.

5 Filtered Images

The noisy images were filtered not only using the algorithms implemented in the section above, but also using known filters, such as the Gaussian, Wiener, BM3D, Median, Bilateral, Non-Local Means (NLM), a combination of Variance Stabilizing Transform (VST) and Non-Local Means, Lee, Frost, Log-Wiener and Wiener-Anscombe filters (see, e.g., [3,6,16–18,21,23,24,26,30]), with the aim of comparing the results with NN and SK operators. The images "cameraman.png", "lena.png" and "moon.png" are more representative of the characteristics of the dataset, therefore we take them as sample for this paper. Specifically, the *Cameraman* image contains well-defined edges and structured details, making it suitable for evaluating edge detection and filtering performance. The *Lena* image offers a balanced combination of smooth regions and complex textures, which is particularly useful for assessing compression and denoising algorithms. Finally, the *Moon* image is characterized by irregular natural textures and gradual intensity variations, making it appropriate for testing performance on less structured, real-world patterns.

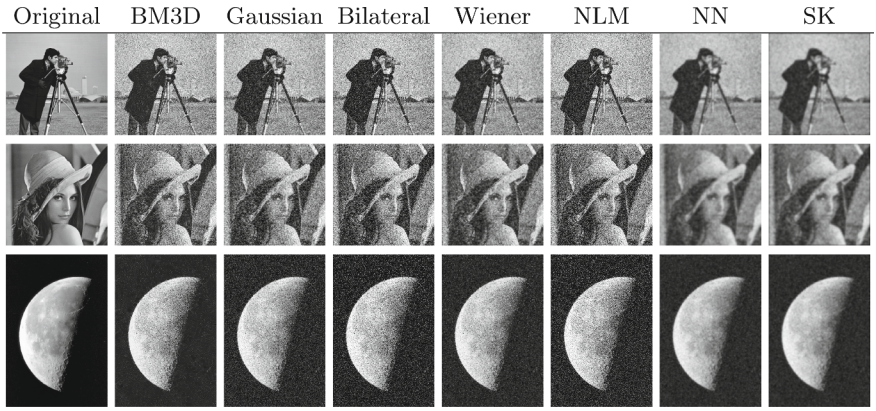


Fig. 2. The test images corrupted by Gaussian noise and then filtered using the study's selected filters for this type of noise. On the left the reference images followed by the filtered images.

The images corrupted by Gaussian noise were filtered using the NN and SK operators, in order to verify their efficiency, and other filters such as the BM3D filter, the Gaussian filter, the Bilateral filter, the Wiener filter and the Non-Local Means (NLM) filter. Figure 2 presents the filtered images corresponding to Gaussian noise, while Fig. 3 presents the filtered images corresponding to Poisson

noise, which were filtered using the NN and SK operators and other employed filters for this type of noise, such as the BM3D, the Median, the Wiener-Anscombe filter, and a combination of Variance Stabilizing Transform (VST) and NLM.

Finally, the filters used for the Speckle noise are the BM3D filter, Lee filter, Frost filter and Log-Wiener filter, along with NN and SK operators. The results are shown in Fig. 4.

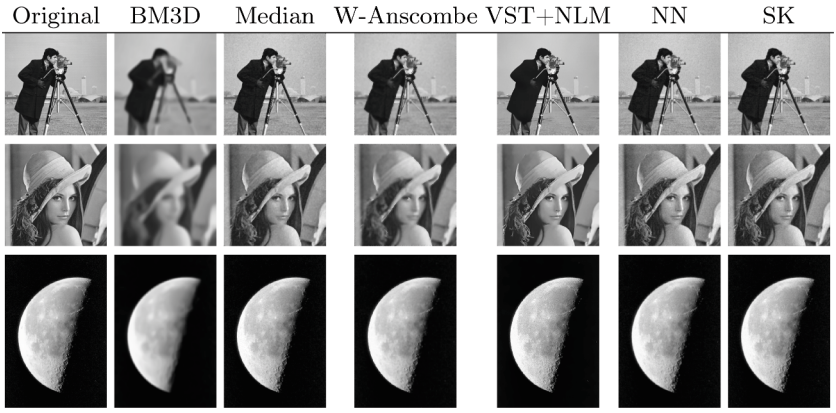


Fig. 3. The test images corrupted by Poisson noise and then filtered using the study's selected filters for this type of noise. On the left the reference images followed by the filtered images.

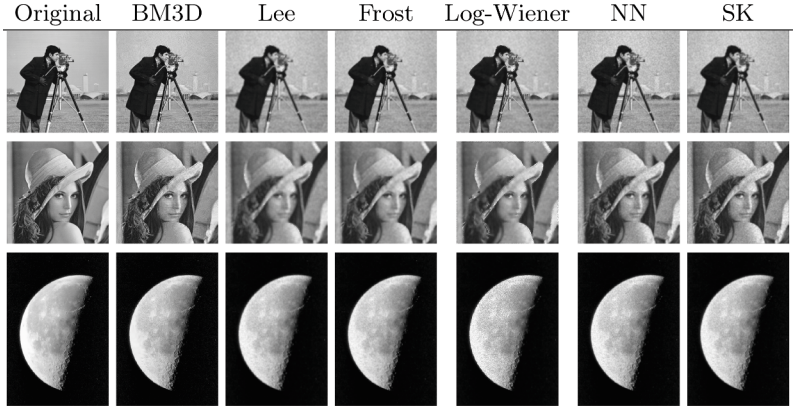


Fig. 4. The test images corrupted by Speckle noise and then filtered using the study's selected filters for this type of noise. On the left the reference images followed by the filtered images.

6 Comparison of the Results

The filtered images are compared with the original images using SSIM, RMSE, FSIM and ꟻLIP indices (see [2,5,29,32]). They were evaluated using MATLAB® formulas, except for the FSIM that was implemented using the code given in [32] and the ꟻLIP index that was calculated using the implementation in [2]. The indices reported for NN and SK algorithms correspond to the best results obtained after the optimization procedure (see Sect. 4). The following tables report the results obtained when filtering images corrupted by Gaussian, Poisson, and Speckle noise. For each noise type, a single table groups the results across all reference images. The columns report the filters used during the elaboration, while the rows are organized by image and quality index.

Table 2. Results obtained filtering images corrupted by Gaussian noise. Each row group corresponds to a reference image; the columns report the quality indices obtained comparing the reference image with each filtered version.

		BM3D	Gaussian	Bilateral	Wiener	NLM	NN	SK
Autumn	↑ **SSIM**	0.2352	0.2584	0.1793	0.3847	0.1637	0.5780	0.5407
	↓ **RMSE**	0.1408	0.1251	0.1741	0.0933	0.1882	0.0764	0.0802
	↑ **FSIM**	0.6333	0.6220	0.5709	0.7118	0.5499	0.7587	0.7527
	↓ **ꟻLIP**	0.2294	0.2259	0.2627	0.2068	0.2791	0.2019	0.2077
Cameraman	↑ **SSIM**	0.2194	0.2361	0.1777	0.3321	0.1655	0.5417	0.5405
	↓ **RMSE**	0.1642	0.1338	0.1876	0.0988	0.2016	0.0874	0.0950
	↑ **FSIM**	0.5794	0.5753	0.5301	0.6624	0.5102	0.7344	0.7141
	↓ **ꟻLIP**	0.2261	0.2311	0.2604	0.2123	0.2764	0.1993	0.2139
Lena	↑ **SSIM**	0.2351	0.2752	0.1950	0.3984	0.1790	0.5967	0.5794
	↓ **RMSE**	0.1625	0.1318	0.1860	0.0955	0.2001	0.0779	0.0791
	↑ **FSIM**	0.5845	0.5859	0.5294	0.6944	0.5068	0.7981	0.7914
	↓ **ꟻLIP**	0.2408	0.2285	0.2729	0.2039	0.2923	0.1988	0.2015
Moon	↑ **SSIM**	0.1591	0.0916	0.0604	0.1546	0.0503	0.3691	0.3910
	↓ **RMSE**	0.1153	0.1243	0.1591	0.1051	0.1743	0.0794	0.0780
	↑ **FSIM**	0.5850	0.4397	0.3942	0.5556	0.3587	0.8366	0.8505
	↓ **ꟻLIP**	0.2368	0.2919	0.3313	0.2578	0.3597	0.2330	0.2289
Onion	↑ **SSIM**	0.1423	0.1877	0.1174	0.3413	0.1047	0.6829	0.6787
	↓ **RMSE**	0.1623	0.1304	0.1860	0.0874	0.2009	0.0534	0.0622
	↑ **FSIM**	0.4857	0.5042	0.4420	0.6398	0.4183	0.8300	0.8162
	↓ **ꟻLIP**	0.2337	0.2161	0.2614	0.1903	0.2804	0.1630	0.1760

The analysis of Table 2 shows that the SSIM index is closest to 1 when using the NN or SK algorithms, regardless of the reference image considered. Similarly, the RMSE values are closer to zero for these two methods, indicating better pixel-wise similarity with respect to the original images. The FSIM index also attains its best values with NN and SK filtering, and the ꟻLIP index confirms this trend, with lower values—i.e., smaller perceptual error—consistently achieved by the same two approaches. These outcomes support the initial hypothesis, confirming the effectiveness of the proposed methods across all tested images. The ꟻLIP index produces an error map that precisely represents the perceived difference between the two images at each pixel in terms of its magnitude. Typically, lower error values are represented by darker or cooler colors (e.g., blue), indicating high similarity, while higher error values correspond to brighter or warmer colors (e.g., yellow or red), highlighting regions where significant discrepancies occur. Therefore, predominantly blue regions in the error map indicate that the filtering method effectively preserves the original image content, whereas the presence of warmer colors reveals areas of degradation or loss of detail.

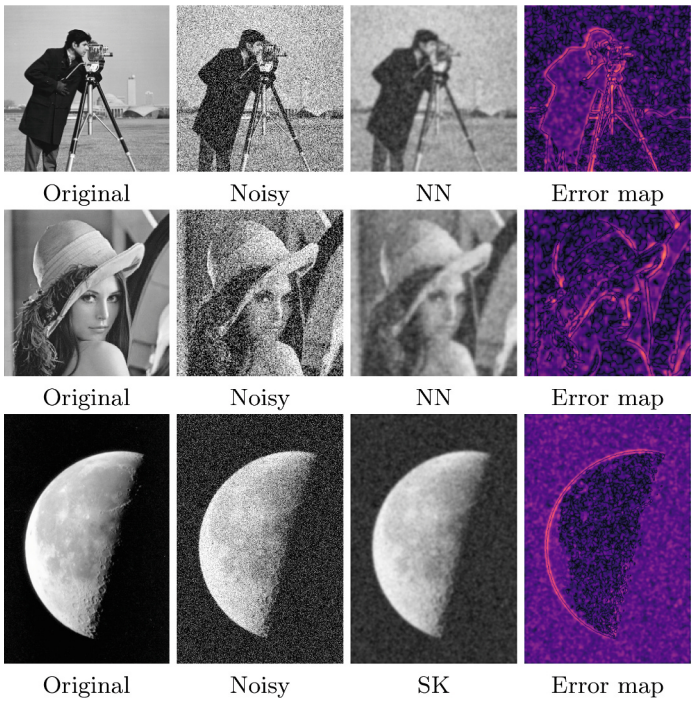


Fig. 5. On the left the reference images, in the center the image with Gaussian noise and the best filtered images, on the right the ꟻLIP error map obtained.

In Fig. 5, the ꟻLIP error maps of the best filtering results are presented for each image considered in the Gaussian noise case.

Table 3. Results obtained filtering images corrupted by Poisson noise. Each row group corresponds to a reference image; the columns report the quality indices obtained comparing the reference image with each filtered version.

		BM3D	Median	W-Anscombe	VST+NLM	NN	**SK**
Autumn	↑ **SSIM**	0.5793	0.7541	0.6965	0.9058	0.8199	0.8029
	↓ **RMSE**	0.0771	0.0437	0.0545	0.0227	0.0337	0.0321
	↑ **FSIM**	0.6585	0.8872	0.8021	0.9410	0.9036	0.9041
	↓ **ꟻLIP**	0.1590	0.0792	0.0969	0.0556	0.0711	0.0707
Cameraman	↑ **SSIM**	0.6396	0.7540	0.7659	0.8955	0.7997	0.7785
	↓ **RMSE**	0.0912	0.0478	0.0481	0.0238	0.0404	0.0374
	↑ **FSIM**	0.6836	0.8959	0.8377	0.9341	0.8951	0.8957
	↓ **ꟻLIP**	0.1492	0.0788	0.0875	0.0511	0.0743	0.0752
Lena	↑ **SSIM**	0.6088	0.8163	0.7541	0.9009	0.8464	0.8447
	↓ **RMSE**	0.0828	0.0378	0.0485	0.0233	0.0304	0.0312
	↑ **FSIM**	0.7404	0.9230	0.8601	0.9392	0.9296	0.9291
	↓ **ꟻLIP**	0.1853	0.0742	0.0982	0.0617	0.0657	0.0676
Moon	↑ **SSIM**	0.8088	0.8971	0.9066	0.8304	0.9251	0.9284
	↓ **RMSE**	0.0291	0.0159	0.0157	0.0257	0.0168	0.0145
	↑ **FSIM**	0.8171	0.9329	0.9197	0.8705	0.9399	0.9442
	↓ **ꟻLIP**	0.0645	0.0332	0.0325	0.0335	0.0336	0.0310
Onion	↑ **SSIM**	0.7357	0.8675	0.8444	0.9185	0.8979	0.8683
	↓ **RMSE**	0.0554	0.0242	0.0316	0.0192	0.0253	0.0278
	↑ **FSIM**	0.8009	0.9340	0.8953	0.9497	0.9303	0.9264
	↓ **ꟻLIP**	0.1467	0.0622	0.0764	0.0595	0.0654	0.0675

Table 3 reveals a more nuanced picture compared to the Gaussian case. The VST+NLM method achieves the best index values across most images and metrics, attaining the highest SSIM, lowest RMSE, highest FSIM, and lowest ꟻLIP scores on Autumn, Cameraman, Lena, and Onion. The NN and SK operators, while not consistently topping the rankings, frequently appear as the second-best performers and remain competitive throughout. On the Moon image, in particular, the SK filter attains the best values for all four indices, with NN ranking second. Notably, the absolute values of the ꟻLIP index across the entire table are significantly lower than those obtained in the Gaussian noise case. The corresponding error maps appear predominantly blue, indicating a high degree of similarity between the filtered and original images when correcting Poisson noise (Fig. 6).

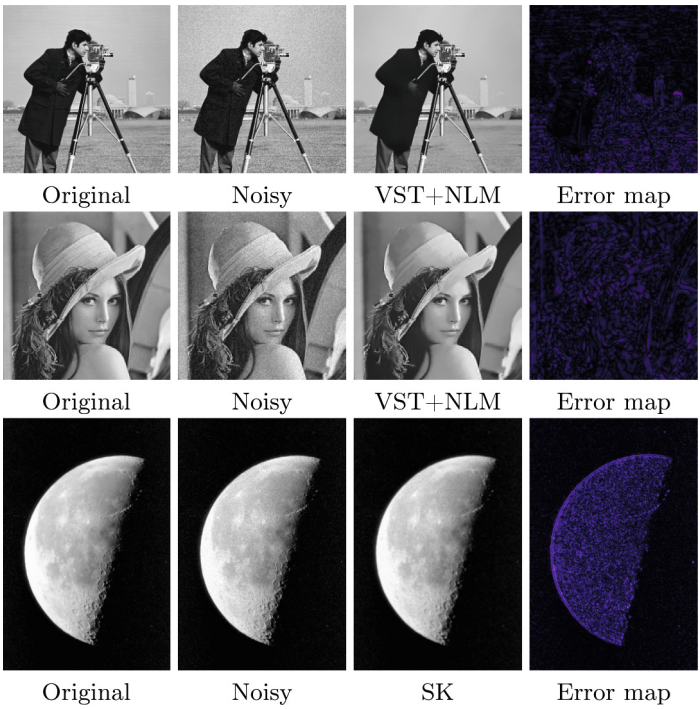


Fig. 6. On the left the reference images, in the center the image with Poisson noise and the best filtered images, on the right the ꟻLIP error map obtained.

As regards Speckle noise, Table 4 shows a more uniform distribution of best results across the competing methods. The BM3D filter achieves the best SSIM and RMSE values on the Autumn image, while the Frost filter ranks first on several other images in terms of SSIM, RMSE, and FSIM. Nevertheless, the NN and SK operators remain consistently close to the top performers, frequently appearing as the second-best method and never falling far behind the leader on any image or metric. This confirms their overall good applicability even in the presence of multiplicative noise. The error maps of the best filtering results are presented in Fig. 7.

Table 4. Results obtained filtering images corrupted by Speckle noise. Each row group corresponds to a reference image; the columns report the quality indices obtained comparing the reference image with each filtered version.

		BM3D	Lee	Frost	Log-Wiener	NN	SK
Autumn	↑ **SSIM**	0.8433	0.6196	0.6975	0.6698	0.6938	0.6883
	↓ **RMSE**	0.0408	0.0655	0.0555	0.0548	0.0525	0.0535
	↑ **FSIM**	0.9152	0.7663	0.8129	0.8156	0.8231	0.8226
	↓ **HLIP**	0.1034	0.1332	0.1231	0.1284	0.1197	0.1194
Cameraman	↑ **SSIM**	0.6535	0.6415	0.7008	0.6889	0.6391	0.6361
	↓ **RMSE**	0.0470	0.0687	0.0563	0.0503	0.0595	0.0613
	↑ **FSIM**	0.8743	0.7689	0.8253	0.8370	0.8031	0.7980
	↓ **HLIP**	0.0854	0.1219	0.1137	0.1085	0.1206	0.1258
Lena	↑ **SSIM**	0.7692	0.7196	0.7765	0.7425	0.7522	0.7430
	↓ **RMSE**	0.0426	0.0555	0.0484	0.0475	0.0494	0.0497
	↑ **FSIM**	0.9043	0.8431	0.8779	0.8752	0.8718	0.8699
	↓ **HLIP**	0.0862	0.1127	0.1043	0.1043	0.1079	0.1074
Moon	↑ **SSIM**	0.8642	0.8648	0.8881	0.7670	0.8755	0.8811
	↓ **RMSE**	0.0281	0.0277	0.0247	0.0468	0.0304	0.0264
	↑ **FSIM**	0.9045	0.8918	0.9118	0.8297	0.8865	0.9000
	↓ **HLIP**	0.0460	0.0525	0.0492	0.0525	0.0570	0.0523
Onion	↑ **SSIM**	0.7491	0.8145	0.8396	0.7761	0.8175	0.7992
	↓ **RMSE**	0.0369	0.0358	0.0323	0.0389	0.0404	0.0422
	↑ **FSIM**	0.8899	0.8807	0.8988	0.8853	0.8733	0.8786
	↓ **HLIP**	0.0871	0.0920	0.0888	0.0923	0.1068	0.1043

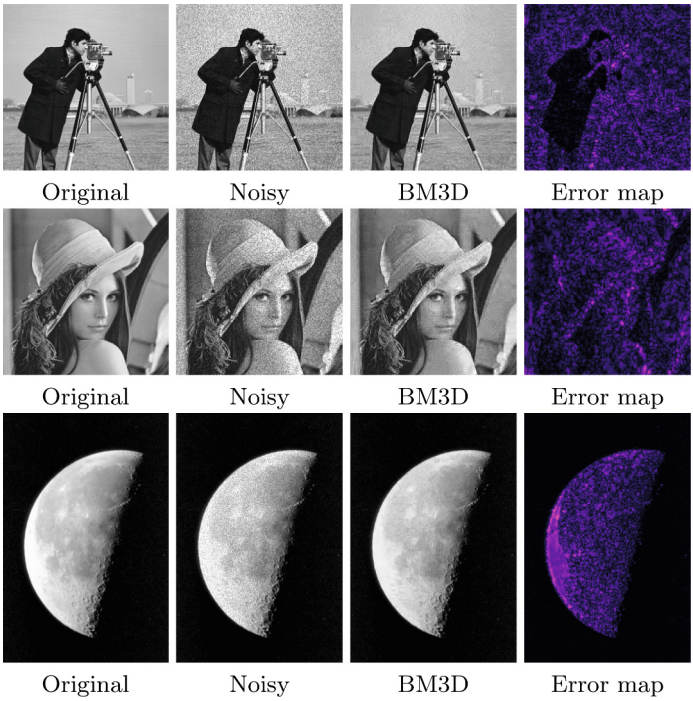


Fig. 7. On the left the reference images, in the center the image with Speckle noise and the best filtered images, on the right the ꟻLIP error map obtained.

Overall, the reconstructions obtained using NN and SK operators proved to be highly competitive across all three noise types considered. They achieved the best results in the Gaussian noise scenario, remained strong second-best alternatives in the Poisson case—where VST+NLM held the edge—and demonstrated solid performance under Speckle noise, where classical filters such as Frost showed a slight advantage. Even in these latter scenarios, the performance gap between operator-based methods and classical filters remains minimal, thus confirming the overall versatility and robustness of the proposed approximation-theoretic framework. These findings suggest that NN and SK operators represent promising alternatives to traditional denoising techniques, offering a more general and flexible approach that is not limited to a specific noise model. Future work may explore their application to other imaging tasks and their integration with hybrid filtering strategies.

Acknowledgments. The authors have been supported within the project "Mathematical Methods and Models for Biomedical Applications" financied by National Recovery and Resilence Plan PNRR-III-C9-2022-I8.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Acu, A.M., Ilina, M., Sofonea, F., Travaglini, A., Vinti, G.: A comparison between neural network and sampling Kantorovich operators in terms of image denoising. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15892, pp. 379–399. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97638-4_24
2. Andersson, P., Nilsson, J., Akenine-Möller, T., Oskarsson, M., Åström, K., Fairchild, M.D.: FLIP: a difference evaluator for alternating images. Proc. ACM Comput. Graph. Interact. Tech. **3**(2), 1–23 (2020)
3. Anscombe, F.J.: The transformation of Poisson, binomial and negative-binomial data. Biometrika **35**(3/4), 246–254 (1948)
4. Bardaro, C., Butzer, P.L., Stens, R.L., Vinti, G.: Kantorovich-type generalized sampling series in the setting of orlicz spaces. Sampl. Theory Signal Image Process. **6**(1), 29–52 (2007)
5. Brunet, D., Vrscay, E.R., Wang, Z.: On the mathematical properties of the structural similarity index. IEEE Trans. Image Process. **21**(4), 1488–1499 (2011)
6. Buades, A., Coll, B., Morel, J.M.: A non-local algorithm for image denoising. In: Proceedings of the IEEE Computer Social Conference on Computing Vision Pattern Recognition (CVPR), vol. 2, pp. 60–65 (2005)
7. Cao, F., Chen, Z.: The approximation operators with sigmoidal functions. Comput. Math. Appl. **58**(4), 758–765 (2009)
8. Cao, F., Chen, Z.: The construction and approximation of a class of neural networks operators with ramp functions. J. Comput. Anal. Appl. **14**(1), 101–112 (2012)
9. Cardaliaguet, P., Euvrard, G.: Approximation of a function and its derivative with a neural network. Neural Netw. **5**(2), 207–220 (1992)
10. Cheang, G.: Approximation with neural networks activated by ramp Sigmoids. J. Approx. Theory **162**, 1450–1465 (2010)
11. Cluni, F., Gusella, V., Vinti, G.: Masonry elastic characteristics assessment by thermographic images. Meccanica **54**(9), 1339–1349 (2019). https://doi.org/10.1007/s11012-019-00982-9
12. Corso, R., Vinti, G.: Mean sampling Kantorovich operators: approximation results and applications. Rev. Real Acad. Cienc. Exactas Fis. Nat. Ser. A-Mat. **120**, 36 (2026). https://doi.org/10.1007/s13398-026-01829-1
13. Costarelli, D., Vinti, G.: Approximation by multivariate generalized sampling kantorovich operators in the setting of orlicz spaces. Boll. Unione Mat. Ital. (9) **IV**, 14–20 (2011)
14. Costarelli, D., Vinti, G.: Convergence for a family of neural network operators in Orlicz spaces. Math. Nachr. **290**(2–3), 226–235 (2017)
15. Cybenko, G.: Approximation by superpositions of a sigmoidal function. Math. Control Sig. Syst. **2**, 303–314 (1989)
16. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. IEEE Trans. Image Process. **16**(8) (2007)
17. Frost, V.S., et al.: A model for radar images and its application to adaptive digital filtering of multiplicative noise. IEEE Trans. Pattern Anal. Mach. Intell. **4**(2), 157–166 (1982)
18. Gonzales, R.C., Woods, R.E.: Digital Image Processing. Addison Wesley (1992)
19. Jaybhay, J., Shastri, R.: A study of Speckle Noise reduction filters. Signal Image Process.: Int. J. (SIPIJ) **6**(3) (2015)

20. Kyurkchiev, N., Markov, S.: Sigmoid functions: some approximation and modelling aspects, some moduli in programming environment mathematica. LAP Lambert Acad. Publ. **4**, 34 (2015)
21. Lee, J.S.: Digital image enhancement and noise filtering by use of local statistics. IEEE Trans. Pattern Anal. Mach. Intell. **2**(2), 165–168 (1980)
22. Longaroni, M., Rinelli, L., Travaglini, A., Vinti, G.: Mathematical models and algorithms for processing retinal OCTA images. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15892, pp. 358–378. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97638-4_23
23. Mäkinen, Y., Azzari, L., Foi, A.: Collaborative filtering of correlated noise: exact transform-domain variance for improved shrinkage and patch matching. IEEE Trans. Image Process. **29** (2020)
24. Makitalo, M., Foi, A.: Optimal inversion of the generalized Anscombe transformation for Poisson-Gaussian noise. IEEE Trans. Image Process. **22**(1), 91–103 (2013)
25. Rinelli, L., Travaglini, A., Vescera, N., Vinti, G.: Sampling type operators versus AI in medical image processing. Discrete Contin. Dyn. Syst. Ser. S (2025). https://doi.org/10.3934/dcdss.2025147
26. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: Proceedings of the 6th International Conference on Computer Vision, pp. 839–846 (1998)
27. Travaglini, A., Vinti, G.: A mathematical model for the analysis of eye fundus images in healthy and diabetic patients. In: Gervasi, O., et al. (eds.) ICCSA 2023. LNCS, vol. 14108, pp. 558–576. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-37117-2_38
28. Travaglini, A., Vinti, G., Scalera, G.B., Scialpi, M.: A large scale analysis for testing a mathematical model for the study of vascular pathologies. Mathematics **11**(8), 1–19 (2023)
29. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)
30. Wiener, N.: Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications. Wiley (1949)
31. Zaabouli, Z., Afraites, L., Laghrib, A., Hadri, A.: Gaussian and Poisson noise identification using non-convex optimization with L^2-norm power constraints. ESAIM Math. Model. Numer. Anal. **59**(3), 1705–1727 (2025). https://doi.org/10.1051/m2an/2025043
32. Zhang, L., Zhang, L., Mou, X., Zhang, D.: FSIM: a feature similarity index for image quality assessment. IEEE Trans. Image Process. **20**(8) (2011)

Multivariate Generalized Gamma-Type Operators: Approximation Results and Applications to Image Processing

Vita Leonessa[1(✉)], Arianna Travaglini[1], and Gianluca Vinti[2]

[1] Dipartimento di Ingegneria, Università degli Studi della Basilicata, Potenza, Italy
{vita.leonessa,arianna.travaglini}@unibas.it
[2] Dipartimento di Matematica e Informatica, Università degli Studi di Perugia, Perugia, Italy
gianluca.vinti@unipg.it

Abstract. In the present paper, we extend a class of integral operators of probabilistic type, which are constructed by means of the generalized Gamma distribution, to the multidimensional setting. In particular, we discuss their approximation properties in weighted continuous function spaces for a broad range of weights, as well as some shape-preserving properties for them. A graphical example is also shown. The paper concludes with applications to image processing, comparing the performance of the new operators, in terms of image denoising, with other well-established denoising techniques.

Keywords: Generalized Gamma operators · weighted approximation · shape-preserving properties · signal and image processing

1 Introduction

In approximation theory, positive linear operators play a central role due to their ability to approximate functions while preserving qualitative properties such as positivity and monotonicity. Classical examples include the Bernstein and Kantorovich operators, which have been extensively studied and generalized in various directions (see, e.g. [3,5,38] and the references therein).

A significant line of research concerns integral operators defined via kernel functions, which are particularly effective at approximating functions on unbounded domains. In this context, operators generated by probability densities have received considerable attention (see, e.g., [3,11,12,20,24,27]).

Among operators of probabilistic type, those associated with Gamma random processes form an important class. Originally introduced by Lupas and Müller in [26], Gamma operators have been thoroughly investigated in the univariate case with respect to their approximation properties, including rates of convergence and Voronovskaya-type theorems (see, e.g., [13,41]). Various generalizations and modifications have also been proposed in recent years, further highlighting their flexibility and effectiveness in approximation processes (see, e.g. [1,10,28]).

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 142–155, 2027.
https://doi.org/10.1007/978-3-032-30530-5_9

Despite the extensive literature on univariate Gamma-type operators, their study in the multidimensional setting remains relatively limited (see [18,21]). This observation provides a natural motivation for extending such operators to functions of several variables, combining the probabilistic structure of Gamma kernels with the analytical framework of multivariate approximation theory.

In particular, in this paper we introduce a multidimensional version of the Gamma-type integral operators previously studied in the univariate case in [12]. Their definition (see Sect. 2) is as follows:

$$G_{n,\underline{\nu}}(f)(\underline{x}) = \frac{n^d}{\prod_{i=1}^{d} \Gamma(\nu_i)} \int_{\mathbb{R}_+^d} \underline{u}^{n\underline{\nu}-\underline{1}} e^{-\sum_{i=1}^{d} u_i^n} f(\underline{x} \odot \underline{u}) d\underline{u}, \tag{1}$$

where $\underline{\nu} = (\nu_1, \ldots, \nu_d) \in \mathbb{R}^d$ with $\nu_i \geq 1$ $(i = 1, \ldots, d)$, $n \geq 1$, and $\underline{x} \in \mathbb{R}_+^d = [0, +\infty)^d$. The symbol $\odot$ denotes the Hadamard product, and f is a real-valued function on $\mathbb{R}_+^d$ such that the integral on the right-hand side of (1) is finite.

In Sect. 3, we investigate the approximation properties of operators $G_{n,\underline{\nu}}$ in weighted continuous function spaces and establish some of their shape-preserving properties. Using a graphical example, we illustrate the behavior of such operators when approximating a test function for different values of n. In addition, in Sect. 4 we explore possible applications in image processing and denoising (for similar studies see, e.g., [2,6,7,25,33]). Specifically, we propose the multivariate generalized Gamma-type operators as a structure-preserving smoothing tool that maintains qualitative features of the original image, such as monotonicity, while reducing noise. This property distinguishes the proposed approach from classical smoothing techniques (Gaussian convolution, bilateral filtering, anisotropic diffusion, Wiener filtering), which provide no explicit guarantee that output images respect structural constraints. Applications include Digital Elevation Models (DEMs) in hydrological analysis, diffusion–weighted MRI, HDR tone mapping, infrared thermography, and aerodynamic pressure mapping–domains, where preserving structural coherence is important.

In this perspective, we compare the performance of operators $G_{n,\underline{\nu}}$, in terms of image denoising, with classical methods, using standard metrics (PSNR and SSIM) and introduce a novel *Monotonicity Violation Index* (MVI) to quantify the preservation of structural properties. Our numerical experiments demonstrate that the operators $G_{n,\underline{\nu}}$ outperforms classical smoothing methods in terms of both denoising quality and structural preservation, establishing it as a theoretically grounded and practically effective tool for applications requiring certified structural fidelity.

2 Generalized Gamma-Type Operators

From now on, let $d \geq 1$ be a fixed positive integer and denote by $\mathbb{R}^d$ the Euclidean space of dimension d, endowed with the usual norm $\|\cdot\|_2$. If $\underline{x}, \underline{y} \in \mathbb{R}^d$, the inequality $\underline{x} \leq \underline{y}$ means that $x_i \leq y_i$ for every $i = 1, \ldots, d$. Moreover, we use the symbols $\underline{0} = (0, \ldots, 0)$ and $\underline{1} = (1, \ldots, 1)$. Furthermore, for $\underline{\nu} = (\nu_1, \ldots, \nu_d) \in \mathbb{R}^d$

and $\underline{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$ with $\underline{x} \geq \underline{0}$, set $\underline{x}^{\underline{\nu}} := \prod_{i=1}^{d} x_i^{\nu_i}$. If $\underline{\nu} \geq \underline{0}$ and $\underline{x} \geq \underline{0}$, then $\underline{x}^{\underline{\nu}}$ is defined as above, assuming that $\underline{0}^{\underline{0}} := 1$.

We are interested in working with real valued functions f defined on $\mathbb{R}_+^d := [0, +\infty)^d$ (in short, $f \in F(\mathbb{R}_+^d)$) belonging to different function spaces. To this end, we fix the following notations used throughout the paper.

First of all, the symbol $C(\mathbb{R}_+^d)$ stands for the space of all continuous real-valued functions on $\mathbb{R}_+^d$. By $C_b(\mathbb{R}_+^d)$ and $C_0(\mathbb{R}_+^d)$ we denote the Banach spaces of all continuous real-valued functions on $\mathbb{R}_+^d$ which are bounded and which vanish at infinity, respectively, endowed with the uniform norm $\|\cdot\|_\infty$ and the pointwise ordering.

Moreover, let w be a weight function on $\mathbb{R}_+^d$, that is $w \in C_b(\mathbb{R}_+^d)$ and $w(x) > 0$ for all $\underline{x} \in \mathbb{R}_+^d$. We consider the space $C_b^w(\mathbb{R}_+^d)$ (resp., $C_0^w(\mathbb{R}_+^d)$) consisting of all those continuous functions f such that $wf \in C_b(\mathbb{R}_+^d)$ (resp., $wf \in C_0(\mathbb{R}_+^d)$). We recall that $C_b^w(\mathbb{R}_+^d)$ (resp., $C_0^w(\mathbb{R}_+^d)$) is a Banach space with respect to the weighted norm $\|\cdot\|_w$ defined as $\|f\|_w = \|wf\|_\infty$ $(f \in C_b^w(\mathbb{R}_+^d))$.

In order to introduce the multivariate operators under study, let us fix $\underline{\nu} \in \mathbb{R}^d$ with $\underline{\nu} \geq \underline{1}$. For every $n \geq 1$, we define the following n-th multivariate operator by setting

$$G_{n,\underline{\nu}}(f)(\underline{x}) = \frac{n^d}{\prod_{i=1}^{d} \Gamma(\nu_i)} \int_{\mathbb{R}_+^d} \underline{u}^{n\underline{\nu}-\underline{1}} e^{-\sum_{i=1}^{d} u_i^n} f(\underline{x} \odot \underline{u}) d\underline{u}, \tag{2}$$

where $\underline{x} \in \mathbb{R}_+^d$, the symbol $\odot$ denotes the Hadamard (componentwise) product, that is $\underline{x} \odot \underline{u} = (x_1 u_1, \ldots, x_d u_d)$, and f is a real-valued function on $\mathbb{R}_+^d$ such that the integral on the right-hand side of (2) is finite.

Note that, for $d = 1$, definition (2) reduces to the unidimensional operators

$$G_{n,\nu}(f)(x) = \frac{n}{\Gamma(\nu)} \int_0^{+\infty} u^{n\nu-1} e^{-u^n} f(xu) du, \quad x \geq 0 \tag{3}$$

which, by a change of variable, turn into the following ones

$$G_{n,\nu}(f)(x) = \frac{n}{x^{n\nu}\Gamma(\nu)} \int_0^{+\infty} t^{n\nu-1} e^{-\left(\frac{t}{x}\right)^n} f(t) dt, \quad x > 0. \tag{4}$$

Operators (4) have been introduced and studied in [12] for suitable functions defined on $(0, +\infty)$. Such operators are called *generalized Gamma-type operators* since they are related to the generalized Gamma distribution (see [35]), whose probability density function is defined by

$$g_{n,x,\nu}(t) = \begin{cases} \frac{n}{x^{n\nu}\Gamma(\nu)} t^{n\nu-1} e^{-\left(\frac{t}{x}\right)^n} & t > 0, \\ 0 & t \leq 0, \end{cases} \tag{5}$$

with $x > 0, n \in \mathbb{N}, n \geq 1$, and $\nu \in \mathbb{R}, \nu \geq 1$. We want to point out that in [12, Proposition 3.2] it was proved that, for every $n \geq 1$, if there exists $\lim_{x \to 0^+} f(x) \in \mathbb{R}$, then $\lim_{x \to 0^+} G_{n,\nu}(f)(x) = \lim_{x \to 0^+} f(x)$.

Therefore, if f is right-continuous at $x = 0$, we can define $G_{n,\nu}(f)(0) := f(0)$, thereby extending the definition of operators (4) to the interval $[0, +\infty)$, as in definition (3). This allows us to review the operators defined in (2) as the multidimensional version of the generalized Gamma-type operators studied in [12]. For this reason, we refer to them as *multivariate generalized Gamma-type operators*. On the other hand, such operators can be considered as operators associated with a vector of independent random variables, each of which has a Gamma-type distribution as in (5):

$$G_{n,\underline{\nu}}(f)(\underline{x}) = \int_{\mathbb{R}_+^d} \prod_{i=1}^{d} g_{n,x_i,\nu_i}(t_i) f(\underline{t}) d\underline{t}. \tag{6}$$

The expression (6) also shows that

$$\text{if } f(\underline{x}) = \prod_{i=1}^{d} f_i(x_i), \text{ then } G_{n,\underline{\nu}}(f)(\underline{x}) = \prod_{i=1}^{d} G_{n,\nu_i}(f_i)(x_i), \tag{7}$$

provided that $G_{n,\nu_i}(f_i)$ is well-defined.

Following the same arguments used in [12] for $d = 1$, the operators (2) are well-defined on the space $E_{\underline{\nu}}$ of all functions $f : \mathbb{R}_+^d \to \mathbb{R}$ such that

$$\int_{\mathbb{R}_+^d} h_{\underline{\nu}}(\underline{u}) |f(\underline{a} \odot \underline{u})| \, d\underline{u} < +\infty \tag{8}$$

for any $\underline{a} \in \mathbb{R}_+^d$, where $h_{\underline{\nu}} : \mathbb{R}_+^d \to \mathbb{R}$ is the integrable function on $\mathbb{R}_+^d$ defined by $h_{\underline{\nu}}(\underline{u}) = \prod_{i=1}^{d} h_{\nu_i}(u_i)$ with

$$h_{\nu_i}(u_i) = \begin{cases} u_i^{\nu_i - 1} e^{-u_i} & 0 \le u_i \le 1 \text{ or } u_i \ge \nu_i, \\ \nu_i^{\nu_i} u_i^{-1} e^{-\nu_i} & 1 < u_i < \nu_i. \end{cases} \tag{9}$$

Hence, we can consider, for every $n \ge 1$, $G_{n,\underline{\nu}} : E_{\underline{\nu}} \to F(\mathbb{R}_+^d)$. Moreover, the operators $G_{n,\underline{\nu}}$ are positive, they are well-defined on $C_b(\mathbb{R}_+^d)$ and, in particular, $\|G_{n,\underline{\nu}}(f)\|_\infty \le \|f\|_\infty$, that is the operators $G_{n,\underline{\nu}}$ are positive linear contractions on $C_b(\mathbb{R}_+^d)$ for every $n \ge 1$.

Example 1. For instance, if $\underline{m} \in \mathbb{N}^d$, the function $f(\underline{x}) = \underline{x}^{\underline{m}}$ $(\underline{x} \in \mathbb{R}_+^d)$ satisfies (8) for every $\underline{\nu} \ge \underline{1}$. This implies that the polynomials on $\mathbb{R}_+^d$ belong to $E_{\underline{\nu}}$ for every $\underline{\nu} \ge \underline{1}$.

Another example is given by the function $f(\underline{x}) = e^{\gamma(\underline{x})}$ $(\underline{x} \in \mathbb{R}_+^d)$, where γ is a positive function such that $\gamma(\underline{x}) = O\left(\sqrt{\prod_{i=1}^{d} x_i}\right)$.

Finally, functions of the type $f(\underline{x}) = \prod_{i=1}^{d} x_i^{\alpha_i} \ln x_i$ $(\alpha_i \ge 0, \underline{x} > \underline{0})$, $f(\underline{0}) = 0$, fulfill condition (8), too.

In what follows, we need the expressions of such operators in the i-th coordinate projections, that is the functions $\mathrm{pr}_i : \mathbb{R}_+^d \to \mathbb{R}_+$, $i = 1, \ldots, d$, defined by $\mathrm{pr}_i(\underline{x}) = x_i$ $(\underline{x} \in \mathbb{R}_+^d)$.

Lemma 1. *For every $n \geq 1$, $h, k \geq 0$, $\underline{x} \in \mathbb{R}_+^d$, $\underline{\nu} \geq \underline{1}$, we have*

$$G_{n,\underline{\nu}}(pr_i{}^h)(\underline{x}) = \frac{\Gamma(\nu_i + h/n)}{\Gamma(\nu_i)} x_i^h \tag{10}$$

and

$$G_{n,\underline{\nu}}(pr_i{}^h pr_j{}^k)(\underline{x}) = \frac{\Gamma(\nu_i + h/n)}{\Gamma(\nu_i)} \frac{\Gamma(\nu_j + k/n)}{\Gamma(\nu_j)} x_i^h x_j^k \tag{11}$$

where $i, j = 1, \ldots, d$, $i \neq j$. In particular,

$$G_{n,\underline{\nu}}(pr_i)(\underline{x}) = \frac{\Gamma(\nu_i + 1/n)}{\Gamma(\nu_i)} x_i, \tag{12}$$

$$G_{n,\underline{\nu}}(pr_i^2)(\underline{x}) = \frac{\Gamma(\nu_i + 2/n)}{\Gamma(\nu_i)} x_i^2. \tag{13}$$

Moreover,

$$G_{n,\underline{\nu}}\left(\sum_{i=1}^d pr_i^2\right)(\underline{x}) = \sum_{i=1}^d \frac{\Gamma(\nu_i + 2/n)}{\Gamma(\nu_i)} x_i^2. \tag{14}$$

Proof. Formulas (10) and (11) follow directly from the definition (2) (or, equivalently, from (6)) of operators $G_{n,\underline{\nu}}$, together with [12, Lemma 2.3], since

$$G_{n,\underline{\nu}}(pr_i{}^h)(\underline{x}) = G_{n,\nu_i}(e_h)(x_i)$$

and

$$G_{n,\underline{\nu}}(pr_i{}^h pr_j{}^k)(\underline{x}) = G_{n,\nu_i}(e_h)(x_i) G_{n,\nu_j}(e_k)(x_k),$$

e_h being the univariate function $e_h(t) = t^h$ $(t \geq 0)$. Formulae (12)-(14) can be derived from (10).

3 Approximation and Shape-Preserving Properties

This section is devoted to the approximation and shape-preserving properties of the multivariate operators $G_{n,\underline{\nu}}$.

First we note that property (7) implies that, for every $n, m \geq 1$, the operators (2) map polynomials on $\mathbb{R}_+^d$ of total degree m into polynomials on $\mathbb{R}_+^d$ of total degree m (see [12, Proposition 2.5]). In particular, the operators $G_{n,\underline{\nu}}$ fix the constant functions.

It is easy to show that operators $G_{n,\underline{\nu}}$ also share the following qualitative properties. In what follows, $\mathrm{Lip}(M, 1)$ stands for the usual space of all Lipschitz continuous functions on $\mathbb{R}_+^d$ with Lipschitz constant M. Note that $\mathrm{Lip}(M, 1) \subset E_{\underline{\nu}}$.

Proposition 1. *Fix $\underline{\nu} \geq \underline{1}$; the following properties are satisfied:*

(a) If $f \in E_{\underline{\nu}}$ is an increasing (decreasing) function with respect to each variable, then, for every $n \geq 1$, $G_{n,\underline{\nu}}(f)$ is increasing (decreasing) with respect to each variable.

(b) If $f \in E_{\underline{\nu}}$ is a convex function with respect to each variable, then, for every $n \geq 1$, $G_{n,\underline{\nu}}(f)$ is convex with respect to each variable.
(c) If $f \in \mathrm{Lip}(M,1)$, then, for every $n \geq 1$, $G_{n,\underline{\nu}}(f) \in \mathrm{Lip}\left(\prod_{i=1}^{d} \frac{\Gamma\left(\nu_i+\frac{1}{n}\right)}{\Gamma(\nu_i)} M, 1\right)$.

Proof. Statement (a) and (b) follow directly from (2). For statement (c),

$$\begin{aligned}
&|G_{n,\underline{\nu}}(f)(\underline{x}) - G_{n,\underline{\nu}}(f)(\underline{y})| \\
&\leq \frac{n^d}{\prod_{i=1}^{d} \Gamma(\nu_i)} \int_{\mathbb{R}_+^d} \underline{u}^{n\underline{\nu}-\underline{1}} e^{-\sum_{i=1}^{d} u_i^n} |f(\underline{x} \odot \underline{u}) - f(\underline{y} \odot \underline{u})| \, d\underline{u} \\
&\leq M \|\underline{x} - \underline{y}\|_2 \frac{n^d}{\prod_{i=1}^{d} \Gamma(\nu_i)} \int_{\mathbb{R}_+^d} \underline{u}^{n\underline{\nu}-\underline{1}} e^{-\sum_{i=1}^{d} u_i^n} \|\underline{u}\|_2 \, d\underline{u} \\
&\leq M \|\underline{x} - \underline{y}\|_2 \frac{n^d}{\prod_{i=1}^{d} \Gamma(\nu_i)} \int_{\mathbb{R}_+^d} \underline{u}^{n\underline{\nu}-\underline{1}} e^{-\sum_{i=1}^{d} u_i^n} \sum_{i=1}^{d} u_i \, d\underline{u} \\
&\leq M \|\underline{x} - \underline{y}\|_2 \prod_{i=1}^{d} \frac{\Gamma\left(\nu_i + \frac{1}{n}\right)}{\Gamma(\nu_i)}.
\end{aligned}$$

In order to discuss the approximation properties of operators in weighted continuous function spaces, let us fix $\underline{\nu} \geq \underline{1}$ and a weight w on $\mathbb{R}_+^d$ such that, for every compact set $K \subset \mathbb{R}_+^d$, there exists $g_{\underline{\nu}} \in L^1(\mathbb{R}_+^d)$ satisfying

$$\frac{h_{\underline{\nu}}(\underline{u})}{w(\underline{a} \odot \underline{u})} \leq g_{\underline{\nu}}(\underline{u}) \quad \text{for every } u \in \mathbb{R}_+^d, a \in K. \tag{15}$$

This implies that $C_b^w(\mathbb{R}_+^d) \subset E_{\underline{\nu}}$ for every $\underline{\nu} \geq \underline{1}$ and, as a consequence, the operators (2) are well-defined also on $C_b^w(\mathbb{R}_+^d)$.

Examples 1. *A particular case of multidimensional weights satisfying condition (15) is given by*

$$w : \mathbb{R}_+^d \to \mathbb{R}, \quad w(\underline{u}) = \prod_{i=1}^{d} w_i(a_i u_i), \tag{16}$$

where, for every $i = 1, \ldots, d$, w_i is a strictly positive bounded function on $[0, +\infty)$ such that, for every compact interval $J \subset [0, +\infty)$, there exists $g_{\nu_i} \in L^1(0, +\infty)$ such that

$$\frac{h_{\nu_i}(u)}{w_i(a_i u_i)} \leq g_{\nu_i}(u) \quad \text{for every } u_i \geq 0, a_i \in J \tag{17}$$

(see (9)). Below we list some examples of univariate weights satisfying (17) which can be used to construct the multivariate weight (16).

1. *Let $w \in C_b([0, +\infty))$ be a weight such that $\inf_{x \geq 0} w(x) > 0$. Hence w satisfies (17). In particular, for every $m \geq 0$, the polynomial weight $w(x) = (1+x^m)^{-1}$ $(x \geq 0)$ satisfies (17).*

2. *Consider $m \geq 1$ and a positive function γ satisfying $\gamma(x) = O(x^{\delta})$, $0 < \delta < 1$. Then $w(x) = e^{-m\gamma(x)}$ $(x \geq 0)$ fulfills* (17).
3. *If $m \in \mathbb{N}$ and $0 < \delta < 1$, then $w(x) = x^{\delta}(1 + x^{m})^{-1}$ $(x \geq 0)$ satisfies* (17). *In all the above examples, condition* (17) *is satisfied as an equality.*

We continue our study with the behavior of the operators on some continuous function spaces.

Proposition 2. *Let w a weight satisfying* (15) *and suppose that, for every $n \geq 1$ and $\underline{\nu} \geq \underline{1}$,*

$$G_{n,\underline{\nu}}(w^{-1}) \in C_b^w(\mathbb{R}_+^d). \tag{18}$$

Then, for every $n \geq 1$ and $\underline{\nu} \geq 1$, the following statements hold:

(a) For every $f \in C_b^w(\mathbb{R}_+^d)$ and $n \geq 1$, $G_{n,\underline{\nu}}(f)$ is continuous on $\mathbb{R}_+^d$. Moreover, $G_{n,\underline{\nu}}(C_b^w(\mathbb{R}_+^d)) \subset C_b^w(\mathbb{R}_+^d)$ and $\|G_{n,\underline{\nu}}\|_{C_b^w(\mathbb{R}_+^d)} = \|G_{n,\underline{\nu}}(w^{-1})\|_w$.
(b) $G_{n,\underline{\nu}}(C_0^w(\mathbb{R}_+^d)) \subset C_0^w(\mathbb{R}_+^d)$ and $\|G_{n,\underline{\nu}}\|_{C_0^w(\mathbb{R}_+^d)} \leq \|G_{n,\underline{\nu}}(w^{-1})\|_w$.

Proof. Let $f \in C_b^w(\mathbb{R}_+^d)$, $n \geq 1$, and $\underline{x} \in \mathbb{R}+^d$. Consider a sequence $(\underline{x}_p)_{p\geq 1}$ in $\mathbb{R}_+^d$ such that $\lim_{p\to+\infty} \underline{x}_p = \underline{x}$. Then,

$$\lim_{p\to+\infty} \underline{u}^{n\underline{\nu}-\underline{1}}\, e^{-\sum_{i=1}^d u_i^n}\, f(\underline{x}_p \odot \underline{u}) = \underline{u}^{n\underline{\nu}-\underline{1}}\, e^{-\sum_{i=1}^d u_i^n}\, f(\underline{x} \odot \underline{u}),$$

for all $\underline{u} \geq \underline{0}$. Now, fix a compact set K of $\mathbb{R}_+^d$ such that $(\underline{x}_p)_{p\geq 1} \subset K$; then there exists $g_{\underline{\nu}} \in L^1(\mathbb{R}_+^d)$ such that

$$\left|\underline{u}^{n\underline{\nu}-\underline{1}}\, e^{-\sum_{i=1}^d u_i^n}\, f(\underline{x}_p \odot \underline{u})\right| \leq \|f\|_w h_{\underline{\nu}}(\underline{u}) w^{-1}(\underline{x}_p \odot \underline{u}) \leq \|f\|_w g_{\underline{\nu}}(\underline{u}).$$

By dominated convergence, $\lim_{p\to+\infty} G_{n,\underline{\nu}}(f)(\underline{x}_p) = G_{n,\underline{\nu}}(f)(\underline{x})$, that is, $G_{n,\underline{\nu}}(f) \in C(\mathbb{R}_+^d)$. Now, taking into account (18) if $f \in C_b^w(\mathbb{R}_+^d)$, for every $\underline{x} \in \mathbb{R}_+^d$,

$$w(\underline{x}) G_{n,\underline{\nu}}(f)(\underline{x}) \leq \|f\|_w w(\underline{x}) G_{n,\underline{\nu}}(w^{-1})(\underline{x}) \leq \|f\|_w \|G_{n,\underline{\nu}}(w^{-1})\|_w.$$

This yields $G_{n,\underline{\nu}}(f) \in C_b^w(\mathbb{R}_+^d)$ and $\|G_{n,\underline{\nu}}\|_{C_b^w(\mathbb{R}_+^d)} \leq \|G_{n,\underline{\nu}}(w^{-1})\|_w$. On the other hand, $\|G_{n,\underline{\nu}}(w^{-1})\|_w \leq \|G_{n,\underline{\nu}}\|_{C_b^w(\mathbb{R}_+^d)} \|w^{-1}\|_{w_m} = \|G_{n,\underline{\nu}}\|_{C_b^w(\mathbb{R}_+^d)}$, that is statement (a).

To prove statement (b), fix $f \in C_0^w(\mathbb{R}_+^d)$; hence, for every $\varepsilon > 0$, there exists a compact set $K \subset \mathbb{R}_+^d$ such that $|f| \leq \max_K |f| + \varepsilon w^{-1}$. Therefore,

$$|G_{n,\underline{\nu}}(f)| \leq \max_K |f| + \varepsilon G_{n,\underline{\nu}}(w^{-1}),$$

from which it follows that $G_{n,\underline{\nu}}(f) \in C_0^w(\mathbb{R}_+^d)$, too.

We now provide approximation results for the operators $G_{n,\underline{\nu}}$ within the framework of weighted continuous function spaces.

Theorem 2. *Let $w \in C_0(\mathbb{R}^d_+)$ be a weight satisfying (15) and (18). Moreover, assume that*

$$\|\cdot\|^2 w \in C_0(\mathbb{R}^d_+). \tag{19}$$

Then, for every $\underline{\nu} \geq \underline{1}$ and $f \in C_0^w(\mathbb{R}^d_+)$,

$$\lim_{n\to+\infty} G_{n,\underline{\nu}}(f) = f \quad \text{with respect to } \|\cdot\|_w,$$

the convergence being uniform on any compact subset of $\mathbb{R}^d_+$.

In particular, for every $\underline{\nu} \geq \underline{1}$ and $f \in C_b(\mathbb{R}^d_+)$,

$$\lim_{n\to+\infty} G_{n,\underline{\nu}}(f) = f$$

uniformly on any compact subset of $\mathbb{R}^d_+$.

Proof. It is a consequence of formulas (12) and (14), hypothesis (19), and Korovkin's result in [4, Corollary 6.13].

Remark 1. For every $\underline{m} > \underline{1}$, consider the polynomial weight defined by setting

$$w_{\underline{m}}(\underline{x}) = \frac{1}{1 + x_1^{2m_1} + \ldots + x_d^{2m_d}} \quad \text{for any } \underline{x} \in \mathbb{R}^d_+.$$

Then, $w_{\underline{m}} \in C_0(\mathbb{R}^d_+)$ and it satisfies (15), (18), and (19).

3.1 Graphical Representations

In this section, we apply the *multivariate generalized Gamma-type operators* $G_{n,\underline{\nu}}$ defined in (2) in order to approximate the test function

$$f(x_1, x_2) = x_1 x_2 e^{-(x_1+x_2)}, \qquad (x_1, x_2) \in \mathbb{R}^2_+.$$

For the sake of simplicity, in the following graphical experiments, we consider the case

$$\nu_0 = (\nu_1, \nu_2) = (1.4, 1.4).$$

Figure 1 represents an example of the approximation behavior of G_{n,ν_0} for $n = 5, 40$, over the domain $[0, 3] \times [0, 3]$. In the figure, the original function f is displayed by a black dashed mesh, while the approximation $G_{n,\nu_0}(f)$ is represented by a colored surface.

This graphical example confirms that, as n increases, the operators $G_{n,\nu_0}(f)$ approximate the function f, in agreement with the theoretical convergence results. The accuracy of the approximation is further quantified by the Root Mean Square Error (RMSE), computed over an equally spaced grid of 40×40 points over $[0, 3] \times [0, 3]$, for n varying in $\{5, 10, 20, 40\}$, and shown in Table 1.

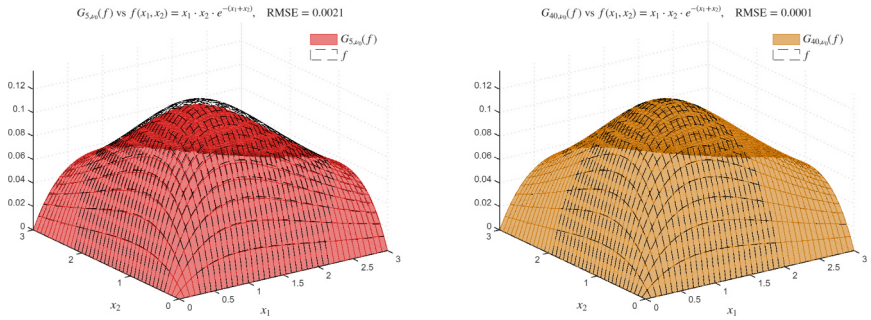


Fig. 1. Approximation of $f(x_1, x_2) = x_1 x_2 e^{-(x_1+x_2)}$ (black dashed mesh) by G_{5,ν_0} (left) and G_{40,ν_0} (right), over $[0,3] \times [0,3]$.

Table 1. RMSE values for the multivariate generalized Gamma-type operators G_{n,ν_0} with $\nu_0 = (1.4, 1.4)$, for the test function f, computed on a 40×40 grid over $[0,3] \times [0,3]$.

RMSE			
$n = 5$	$n = 10$	$n = 20$	$n = 40$
0.0021	0.0006	0.0002	0.0001

4 Structure-Preserving Image Smoothing

Image smoothing is a fundamental pre-processing technique widely employed to suppress noise, which usually occurs in a wide range of applications, such as medical imaging, remote sensing, and engineering fields, causing a degradation of the original signal.

Given a noisy image $f : [0,1]^2 \to [0,1]$, the objective is to produce an approximation $\tilde{f}$ that reduces high-frequency noise while maintaining perceptually relevant structures. Classical methods, such as Gaussian convolution, Bilateral filtering, Anisotropic diffusion or Wiener filter (see, e.g., [29,34,36]) perform well in terms of peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) indices (see [39]). However, they share a common limitation: they provide no explicit guarantee that the output respects structural properties of the input. For instance, if f is known to be monotone (e.g., a digital elevation model, a pressure map, or a radial-density field in MRI), standard smoothing procedures may alter the local ordering of values when applied to noisy and discretized data, potentially introducing artifacts that are not consistent with the underlying physical model. Therefore, in this paper we propose the use of the *multivariate generalized Gamma-type operators* $G_{n,\underline{\nu}}$ as a structure-preserving smoother for such images. In particular we exploit the qualitative properties established in Proposition 1, which, in general, are not shared by any of the classical methods in our comparison. This makes $G_{n,\underline{\nu}}$ the natural choice whenever an application requires certified structural fidelity, not merely statistical denoising quality.

Definition 1. *Let $f : [0,1]^2 \to \mathbb{R}$ be an image satisfying a structural constraint $\mathcal{C}$ (e.g., monotonicity or convexity in a given direction), and let $f_\varepsilon = f + \varepsilon$ be the observed image, where ε denotes noise. A smoother $\mathcal{S}$ is said to be* structure-preserving *with respect to $\mathcal{C}$ if*

$$f \in \mathcal{C} \implies \mathcal{S}(f_\varepsilon) \in \mathcal{C}.$$

Operator $G_{n,\underline{\nu}}$ satisfies the requirement of Definition 1 exactly for monotonicity and convexity, as established in Proposition 1, unlike classical convolutional smoother. Preserving structural properties such as monotonicity and positivity is important in several image processing applications.

For instance, in hydrological analysis of Digital Elevation Models (DEMs), smoothing procedures should not significantly alter the local ordering of elevation values, since this may affect drainage patterns and flow accumulation algorithms (see, e.g., [17,31,40,42]). In practice, standard smoothing techniques applied to noisy and discretized data may introduce small-scale artifacts such as spurious sinks.

Another example is given by diffusion-weighted MRI, in which the apparent diffusion coefficient (ADC) is expected to vary consistently along white matter tracts, reflecting anisotropic diffusion processes (see, e.g., [9,19,22,30]). Smoothing procedures that distort this structural coherence may affect downstream analysis.

Furthermore, in HDR tone mapping, monotonicity of the mapping operator is required to preserve the relative ordering of luminance values, ensuring that brighter scene points correspond to brighter display intensities (see, e.g., [14,15,23,32]).

Similarly, in applications such as infrared thermography and aerodynamic pressure mapping, preserving qualitative structural properties of the data is important for physical interpretability (see, e.g., [8,16,37]).

In this context, the operator $G_{n,\underline{\nu}}$ preserves monotonicity, positivity, and convexity by construction, due to its representation as a convex combination of sampled values. This provides a theoretical guarantee that such structural properties are not violated during the smoothing process.

To quantify the structural fidelity of a smoother, we introduce the following metric.

Definition 2 (Monotonicity Violation Index). *Let $f : \{0,\dots,H-1\} \times \{0,\dots,W-1\} \to [0,1]$ be a discrete image that is non-decreasing in the horizontal direction, i.e.,*

$$f(i,j) \le f(i,j+1) \quad \text{for every } i \in \{0,\dots,H-1\},\ j \in \{0,\dots,W-2\}.$$

For a processed image $\tilde{f}$ defined on the same domain, the ***Monotonicity Violation Index*** *is defined as*

$$\mathrm{MVI}(\tilde{f}) = \frac{1}{H(W-1)} \sum_{i=0}^{H-1} \sum_{j=0}^{W-2} \max\Big(0, \big(f(i,j+1) - f(i,j)\big)\big(\tilde{f}(i,j) - \tilde{f}(i,j+1)\big)\Big).$$

By Proposition 1, $G_{n,\underline{\nu}}$ achieves MVI = 0 on any monotone input, regardless of n and $\underline{\nu}$.

We now compare $G_{n,\underline{\nu}}$ with the classical smoothing methods, namely Gaussian blur, Bilateral filter, Anisotropic Diffusion and Wiener filter, in terms of PSNR, SSIM, and MVI (Definition 2). We report numerical experiments on a synthetic monotone ramp image $f(i,j) = j$, defined on a 256×256 grid with pixel coordinates normalized to $[0,1]^2$, which is strictly increasing in the horizontal direction. Additive Gaussian noise with zero mean and variance $\sigma^2 = 0.05$ is added. For $G_{n,\underline{\nu}}$ we fix $\nu_1 = \nu_2 = 2$ and select $n = 20$.

Table 2. PSNR (dB), SSIM and MVI resulting by the comparison between the original monotone ramp image and the noisy one ($\sigma^2 = 0.05$). Best value in bold.

Method	PSNR (dB)	SSIM	MVI
Noisy input	14.19	0.0286	0.0004
Gaussian blur	23.61	0.3315	0.0001
Bilateral filter	22.88	0.2368	0.0001
Anisotropic diffusion	14.66	0.0338	0.0003
Wiener filter	23.52	0.3983	0.0001
$G_{n,\underline{\nu}}$, $n = 20$	**27.36**	**0.7897**	**0.0000**

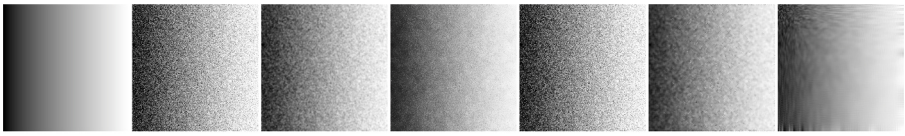


Fig. 2. Monotone ramp image. From left to right: original, noisy ($\sigma^2 = 0.05$), and filtered images.

The results in Table 2 confirm that $G_{n,\underline{\nu}}$ is not only the only method achieving MVI= 0, but also the best performing in terms of PSNR and SSIM, with a significant gap with respect to other classical denoising techniques.

Figure 2 illustrates the visual comparison, showing the original ramp, the noisy version, and the outputs of each method.

Acknowledgments. The paper has been performed within the activities of the Italian GNAMPA-INdAM Group, the UMI Group TAA "Approximation Theory and Applications" and, of the network RITA (Research ITalian network on Approximation). The first and second authors also acknowledge support for the project "Tech4You–Technologies for climate change adaptation and quality of life improvement", Mission

4, Component 2, Investment 1.5, funded by the Next Generation EU—Italian National Recovery and Resilience Plan (NRRP). The second and the third author have been partially supported by University of Perugia, Fondo di Ricerca di Ateneo 2022, Project "MiRA: Mixed Reality and AI Methodologies for Immersive Robotics"

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Acar, T., Mursaleen, M., Deveci, S.N.: Gamma operators reproducing exponential functions. Adv. Diff. Equ. **2020**, 423(2020). https://doi.org/10.1186/s13662-020-02880-x
2. Acu, A.M., Ilina, M., Sofonea, F., Travaglini, A., Vinti, G.: A comparison between neural network and sampling Kantorovich operators in terms of image denoising. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15892, pp. 379–399. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97638-4_24
3. Altomare, F., Campiti, M.: Korovkin-type approximation theory and its applications. de Gruyter Studies in Mathematics, 17. Walter de Gruyter & Co., Berlin (1994)
4. Altomare, F.: Korovkin-type theorems and approximation by positive linear operators. Surv. Approx. Theory **5**, 92–164 (2010). https://surveys-in-approximation-theory.com/papers/13/index.html
5. Altomare, F., Cappelletti Montano, M., Leonessa, V., Raşa, I.: A generalization of Kantorovich operators for convex compact subsets. Banach J. Math. Anal. **11**(3), 591–614 (2017). https://doi.org/10.1215/17358787-2017-0008
6. Angeloni, L., Costarelli, D., Vinti, G.: Convergence in variation for the multidimensional generalized sampling series and applications to smoothing for digital image processing. Ann. Acad. Sci. Fenn. Math. **45**(2), 751–770 (2020). https://doi.org/10.5186/aasfm.2020.4532
7. Asdrubali, F., et al.: Mathematical models for the improvement of detection techniques of industrial noise sources from acoustic images. Math. Methods Appl. Sci. **44**(13), 10448–10459 (2021)
8. Balaras, C.A., Argiriou, A.A.: Infrared thermography for building diagnostics. Energy Build. **34**(2), 171–183 (2002)
9. Basser, P.J., Mattiello, J., LeBihan, D.: MR diffusion tensor spectroscopy and imaging. Biophys. J . **66**(1), 259–267 (1994)
10. Braha, N.L.: Approximation by modified (p, q)-gamma-type operators. J. Inequal. Appl. **2024**, 37 (2024)
11. Cappelletti Montano, M., Leonessa, V.: A note on Weibull operators and their modifications. Positivity **29**, 4 (2025). https://doi.org/10.1007/s11117-024-01096-7
12. Cappelletti Montano, M., Leonessa, V., Travaglini, A.: On the approximation properties of generalized Gamma-type operators in several function spaces. Ann. Funct. Anal. **16**, 59 (2025). https://doi.org/10.1007/s43034-025-00453-1
13. Chen, W.Z., Guo, S.S.: On the rate of convergence of the Gamma operators for functions of bounded variation. Approx. Theory Appl. **1**(5), 85–96 (1985)
14. Drago, F., Myszkowski, K., Annen, T., Chiba, N.: Adaptive logarithmic mapping for displaying high contrast scenes. Comput. Graph. Forum **22**(3), 419–426 (2003)

15. Durand, F., Dorsey, J.: Fast bilateral filtering for the display of high-dynamic-range images. ACM Trans. Graph. **21**(3), 257–266 (2002)
16. Gaussorgues, G., Chomet, S.: Infrared Thermography. Springer, Dordrecht (2012)
17. Hutchinson, M.F.: A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. J. Hydrology **106**(3–4), 211–232 (1989)
18. Izgi, A.: Order of approximation of functions of two variables by new type gamma operators. Gen. Math. **17**(1), 23–32 (2009)
19. Jones, D.K.: The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: a Monte Carlo study. Magn. Reson. Med. **51**(4), 807–815 (2004)
20. Karlin, S., Micchelli, C.A., Rinott, Y.: Multivariate splines: a probabilistic perspective. J. Multivariate Anal. **20**(1), 69–90 (1986)
21. Khan, T., Khan, S.A., Rao, N.: On approximation by bivariate Szász-gamma type hybrid operators. Bull. Math. Anal. Appl. **14**(1), 11–27 (2022)
22. Le Bihan, D., Breton, E., Lallemand, D., Grenier, P., Cabanis, E., Laval-Jeantet, M.: MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. Radiology **161**(2), 401–407 (1986)
23. Ledda, P., Chalmers, A., Troscianko, T., Seetzen, H.: Evaluation of tone mapping operators using a high dynamic range display. ACM Trans. Graph. **24**(3), 640–648 (2005)
24. Leonessa, V., Travaglini, A.: On a sequence of integral operators associated with power Lindley distribution. Carpathian J. Math. **42**(4), 807–825 (2026)
25. Longaroni, M., Rinelli, L., Travaglini, A., Vinti, G.: Mathematical models and algorithms for processing retinal OCTA images. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15892, pp. 358–378. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97638-4_23
26. Lupaş, A., Müller, M.: Approximationseigenschaften der Gammaoperatoren. Math. Z. **98**, 208–226 (1967)
27. Milella, S.: Multivariate approximation by modified Gauss-Weierstrass operators. Rend. Circ. Mat. Palermo (2) **65**(1), 1–15 (2016)
28. Özçelik, R., Kara, E.E., Usta, F., Ansari, K.J.: Approximation properties of a new family of Gamma operators and their applications. Adv. Difference Equ. **2021**(1), 1–13 (2021). https://doi.org/10.1186/s13662-021-03666-5
29. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. **12**(7), 629–639 (1990)
30. Pierpaoli, C., Jezzard, P., Basser, P.J., Barnett, A., Di Chiro, G.: Diffusion tensor MR imaging of the human brain. Radiology **201**(3), 637–648 (1996)
31. Planchon, O., Darboux, F.: A fast, simple and versatile algorithm to fill the depressions of digital elevation models. CATENA **46**(2–3), 159–176 (2001)
32. Reinhard, E., Heidrich, W., Debevec, P., Pattanaik, S., Ward, G., Myszkowski, K.: High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting, 2nd edn. Morgan Kaufmann (2010)
33. Rinelli, L., Travaglini, A., Vescera, N., Vinti, G.: Sampling type operators versus AI in medical image processing. Discrete Contin. Dyn. Syst. Ser. S (2025). https://doi.org/10.3934/dcdss.2025147
34. Rudin, L., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D **60**, 259–268 (1992)
35. Stacy, E.W., Mihram, G.A.: Parameter estimation for a generalized gamma distribution. Technometrics **7**(3), 349–358 (1965)
36. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: Proceedings of the ICCV, pp. 839–846 (1998)

37. Tropea, C., Yarin, A.L., Foss, J.F. (eds.): Springer Handbook of Experimental Fluid Mechanics. Springer, Berlin (2007)
38. Vinti, G., Zampogni, L.: Approximation results for a general class of Kantorovich type operators. Adv. Nonlinear Stud. **14**(4), 991–1011 (2014)
39. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)
40. Wang, L., Liu, H.: An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling. Int. J. Geograph. Inf. Sci. **20**(2), 193–213 (2006)
41. Zeng, X.M.: Approximation properties of gamma operators. J. Math. Anal. Appl. **311**(2), 389–401 (2005). https://doi.org/10.1016/j.jmaa.2005.02.051
42. Zhang, W., Chen, Y., Li, J.: Smoothing of digital elevation models and its effects on terrain analysis. Geomorphology **253**, 66–75 (2016)

Check for updates

Parameters Estimation via Unrolling for the Preconditioned Nested Primal Dual Algorithm

Marco Donatelli and Giuseppe Scarlato(✉)

Università dell'Insubria, 22100 Como, CO, Italy
{marco.donatelli,gscarlato}@uninsubria.it

Abstract. Variational methods are widely used in inverse problems. Through regularization, they are able to incorporate prior knowledge on the solution and, therefore, to achieve good reconstruction quality. However, those methods typically require careful tuning of the parameters involved in the optimization problem and the method itself, which can be time-consuming and computationally expensive. On the other hand, deep learning methods allow for automatic learning through supervised training at the cost of interpretability, generalization, and inclusion of constraints and physical models. Unrolling techniques have emerged as a bridge between these two approaches, combining the interpretability of deterministic methods with the learning capabilities of deep learning to automatically estimate parameters.

In this work, we propose an unrolling architecture for the *Preconditioned Nested Primal Dual* (PNPD) algorithm. This allows to automatically estimate the parameters involved in the optimization problem and in the numerical method, without the need for manual tuning. Numerical results show the effectiveness of the proposed approach in terms of reconstruction quality and parameter estimation.

Keywords: Proximal algorithms · Preconditioning · Unrolling · Deep unfolding · Neural network · Image restoration · Regularization

1 Introduction

We consider a linear inverse problem of the form

$$b^\delta = Au + \eta_\delta = b + \eta_\delta, \tag{1}$$

where $A \in \mathbb{R}^{s\times d}$ is a linear operator, $b^\delta \in \mathbb{R}^s$ is the observed data corrupted by noise $\eta_\delta \in \mathbb{R}^s$, and $u \in \mathbb{R}^d$ is the unknown signal we want to recover. This kind of problem arises in many applications, such as image restoration, computed tomography, and many others [5,9]. In particular, we will focus on the case of image deblurring, where A is a convolution operator, u is the original image, and b^δ is the blurred and noisy observation. The ill-posed nature of the operator

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 156–169, 2027.
https://doi.org/10.1007/978-3-032-30530-5_10

A and the presence of noise makes the problem challenging, and regularization is typically required to achieve a useful reconstruction. A common variational approach to solve Eq. (1) is to consider the following optimization problem

$$\underset{u\in\mathbb{R}^d}{\operatorname{argmin}} \ \frac{1}{2}\|Au - b^\delta\|^2 + h(Wu) = \underset{u\in\mathbb{R}^d}{\operatorname{argmin}} \ f(u) + h(Wu), \tag{2}$$

where, $\|.\|$ denotes the Euclidean norm, $f(u)$ is the data fidelity term, and $h(Wu)$ is the regularization term, with $W \in \mathbb{R}^{d'\times d}$ being a linear operator and $h : \mathbb{R}^{d'} \to \mathbb{R} \cup \{\infty\}$ is a convex and possibly non-smooth function. The balance between the data fidelity and the regularization is typically controlled by a regularization parameter $\lambda > 0$, which here is implicitly included in the definition of h. To solve Eq. (2), since $h \circ W$ is non-smooth, a common approach is to leverage the proximity operator of $g = h \circ W$:

$$\operatorname{prox}_g(x) = \underset{u\in\mathbb{R}^d}{\operatorname{argmin}} \ \frac{1}{2}\|u - x\|^2 + g(u).$$

This leads to the use of proximal gradient methods, which are iterative algorithms that alternate between a gradient step on the smooth term f and a proximal step on the convex term $h \circ W$, specifically

$$\begin{cases} u_{n+\frac{1}{2}} = u_n - \alpha\nabla f(u_n) \\ u_{n+1} = \operatorname{prox}_{\alpha h\circ W}(u_{n+\frac{1}{2}}) \end{cases}$$

where $\alpha > 0$ is the step size. Common examples of proximal algorithms include the *Iterative Shrinkage Thresholding Algorithm* (ISTA) [14], the *Fast Iterative Shrinkage Thresholding Algorithm* (FISTA) [4], the *iPiano* algorithm [21], *VMILAn* [7], and many others (for a review, see, for example, [13]). An issue that can occur with this approach is that the proximity operator of $h \circ W$ might not be known in closed form. This is the case, for example, of the Total Variation (TV) regularization. To address this, inexact proximal algorithms have been proposed, which, instead of computing the proximal step exactly, approximate it by solving a subproblem iteratively [12,22]. In particular, we are interested in the *Nested Primal Dual* (NPD) algorithm [8,10], which approximates $\operatorname{prox}_{\alpha h\circ W}$ via a dual sequence that depends on the proximity operator of the Fenchel conjugate of h, namely $\operatorname{prox}_{\alpha h^*}$, which in many cases is known in closed form, as in the case of TV regularization. Furthermore, by using a preconditioning strategy, the convergence of the NPD algorithm can be improved. This leads to the *Nested Primal Dual Iterated Tikhonov* (NPDIT) method [1] and the *Preconditioned Nested Primal Dual* (PNPD) algorithm [2], which is the focus of this work.

Variational methods, such as the PNPD algorithm, often depend on several parameters, including the regularization parameter λ, the step sizes of the algorithm, and the parameters of the preconditioner. Choosing these parameters is challenging, as they interact in nontrivial ways and can significantly affect both the reconstruction quality and the convergence of the algorithm. These parameters are typically chosen manually, but automatic parameter estimation is possible using heuristic rules, such as the discrepancy principle [23], the L-curve

method [18], and generalized cross-validation [16]. These methods, however, can be computationally expensive and may not always provide optimal results. In recent years, deep learning methods have received a lot of attention in the field of inverse problems. They provide a flexible and powerful tool to learn complex relationships from data. A straightforward approach to leverage neural networks for solving inverse problems is to unroll the iterations of a deterministic algorithm, such as the PNPD algorithm, and to learn the parameters of the algorithm through supervised training [20]. With this approach, the method itself becomes a part of a neural network. In this work, we propose an unrolling architecture for the PNPD algorithm, which allows us to automatically estimate the parameters involved in the optimization problem and in the numerical method, without the need for manual tuning.

This paper is organized as follows: Sect. 2 introduces the PNPD algorithm. Section 3 describes the proposed unrolling architecture for the PNPD algorithm and the loss function used for training. Section 4 presents the numerical results for the image deblurring problem. Finally, Sect. 5 concludes the paper and discusses future work.

2 Preconditioned Nested Primal Dual Algorithm

In this section, we introduce the NPD algorithm and its preconditioned version, the PNPD algorithm. As shown in [2], the NPDIT method proposed in [1] can also be seen as a preconditioned version of NPD. The difference between the two methods resides in how the preconditioning is applied. The PNPD algorithm corresponds to a left preconditioning, while the NPDIT method corresponds to a right preconditioning strategy. In this work, we focus on the PNPD algorithm, but the proposed unrolling architecture can be adapted to the NPDIT method as well.

2.1 Nested Primal-Dual Method (NPD)

The NPD algorithm is an inexact proximal gradient algorithm, and its iterative scheme can be summarized as follows:

$$\begin{cases} \bar{u}_n = u_n + \gamma_n(u_n - u_{n-1}), \\ u_{n+\frac{1}{2}} = \bar{u}_n - \alpha_n \nabla f(\bar{u}_n), \\ u_{n+1} \approx \operatorname{prox}_{\alpha_n h \circ W}(u_{n+\frac{1}{2}}), \end{cases} \tag{3}$$

where $\approx$ indicates an approximation of the proximal-gradient point. Here, $\bar{u}_n$ is usually referred to as the extrapolated point, $\gamma_n \in [0, 1]$ is the extrapolation parameter, and $\alpha_n > 0$ is the step length along the descent direction $-\nabla f(\bar{u}_n)$. For the iterative scheme in Eq. (3), a way to approximate $\operatorname{prox}_{\alpha_n h \circ W}$ is required. This is achieved by computing a dual sequence which, in turn, is used to approximate $\operatorname{prox}_{\alpha_n h \circ W}$. To introduce the dual sequence, we state the following result.

Lemma 1. *[8, Theorem 1] Suppose that $h : \mathbb{R}^{d'} \to \mathbb{R} \cup \{\infty\}$ is a proper convex lower semicontinuous function and $W \in \mathbb{R}^{d' \times d}$ is such that there exists $u_0 \in \mathbb{R}^d$ with $Wu_0 \in \operatorname{relint}(\operatorname{dom}(h))$. Given $a \in \mathbb{R}^d$, $\alpha > 0$, $0 < \beta < 2/\|W\|^2$, and $v^0 \in \mathbb{R}^{d'}$, define the sequence*

$$v^{k+1} = \operatorname{prox}_{\beta\alpha^{-1}h^*}(v^k + \beta\alpha^{-1}W(a - \alpha W^T v^k)), \quad \forall k \geq 0. \tag{4}$$

Then there exists $\hat{v} \in \mathbb{R}^{d'}$ such that:

(i) $\lim_{k\to\infty} v^k = \hat{v}$;
(ii) $\operatorname{prox}_{\alpha h \circ W}(a) = a - \alpha W^T \hat{v}$.

Note that by combining items (i) and (ii) of Lemma 1 with Eq. (4), it holds

$$\operatorname{prox}_{\alpha h\circ W}(a) = \lim_{k\to\infty} a - \alpha W^T \operatorname{prox}_{\beta\alpha^{-1}h^*}(v^k + \beta\alpha^{-1}W(a - \alpha W^T v^k)).$$

Using this relationship, it is possible to compute an approximation of $\operatorname{prox}_{\alpha h\circ W}(a)$ by a finite number of steps of the dual sequence (Eq. (4)), provided that the operator $\operatorname{prox}_{\beta\alpha^{-1}h^*}$ and the matrix-vector product with the linear operators W and W^T are easily computable. Let $k_{\max}$ be the maximum number of inner iterations used to compute the dual sequence Eq. (4), and L be the Lipschitz constant of the gradient of f, the resulting NPD algorithm is summarized in Algorithm 1.

Algorithm 1. NPD

Choose $u_0 \in \mathbb{R}^d$, $0 < \alpha < \frac{2}{L}$, $0 < \beta < \frac{2}{\|W\|^2}$, $k_{\max} \in \mathbb{N}$, $\{\gamma_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}_{\geq 0}$.
$u_{-1} = u_0$
$v_0^0 = 0$
for $n = 0, 1, \ldots$ **do**
  $\bar{u}_n = u_n + \gamma_n(u_n - u_{n-1})$
  $u_{n+\frac{1}{2}} = \bar{u}_n - \alpha\nabla f(\bar{u}_n)$
  for $k = 0, 1, \ldots, k_{\max} - 1$ **do**
    $u_n^k = u_{n+\frac{1}{2}} - \alpha W^T v_n^k$
    $v_n^{k+1} = \operatorname{prox}_{\beta\alpha^{-1}h^*}(v_n^k + \beta\alpha^{-1}Wu_n^k)$
  end for
  $v_{n+1}^0 = v_n^{k_{\max}}$
  $u_n^{k_{\max}} = u_{n+\frac{1}{2}} - \alpha W^T v_n^{k_{\max}}$
  $u_{n+1} = \frac{1}{k_{\max}} \sum_{j=1}^{k_{\max}} u_n^j$
end for

2.2 Preconditioned NPD (PNPD)

The main idea of PNPD is to accelerate the convergence of NPD by replacing the gradient descent step inside the proximal evaluation with a preconditioned

iteration. This is achieved by modifying the data fidelity norm with the one induced by an appropriate positive definite preconditioner $S \in \mathbb{R}^{s\times s}$.

Similarly to Eq. (2), given $S \in \mathbb{R}^{s\times s}$ symmetric positive definite (s.p.d.), we consider the optimization problem

$$\operatorname*{argmin}_{u\in\mathbb{R}^d} f_S(u) + h(Wu) = \operatorname*{argmin}_{u\in\mathbb{R}^d} \frac{1}{2}\|Au - b^\delta\|^2_{S^{-1}} + h(Wu). \tag{5}$$

In the data fidelity term

$$f_S(u) = \frac{1}{2}\|S^{-\frac{1}{2}}(Au - b^\delta)\|^2, \tag{6}$$

the linear operator $S^{\frac{1}{2}}$ can be interpreted as a left preconditioner for the linear system $Ax = b^\delta$. If we further assume that there exists $P \in \mathbb{R}^{d\times d}$ s.p.d. such that

$$P^{-1}A^T = A^T S^{-1}, \tag{7}$$

then it holds

$$\nabla f_S(u) = A^T S^{-1}(Au - b^\delta) = P^{-1}\nabla f(u), \tag{8}$$

where $f(u) = \frac{1}{2}\|Au - b^\delta\|^2$ is as in Eq. (2).

Therefore, applying the NPD algorithm to problem (5), under the Assumption (7), we obtain the iterative scheme of PNPD, defined as

$$\begin{cases} \bar{u}_n = u_n + \gamma_n(u_n - u_{n-1}) \\ u_{n+\frac{1}{2}} = \bar{u}_n - \alpha_n P^{-1}\nabla f(\bar{u}_n) \\ u_{n+1} \approx \operatorname{prox}_{\alpha_n h\circ W}(u_{n+\frac{1}{2}}). \end{cases} \tag{9}$$

A reasonable choice for the preconditioner P and the associated matrix S, that satisfies the condition (7), is

$$P = A^T A + \nu I, \qquad S = AA^T + \nu I, \tag{10}$$

with $\nu > 0$. This is inspired by the iterated Tikhonov method, which coincides with the Levenberg–Marquardt method when applied to linear problems [15].

Algorithm 2 reports the pseudocode of the PNPD method, where we denote by L_S the Lipschitz constant of ∇f_S, which depends on the choice of the preconditioner S, i.e., P, due to (7). In particular, thanks to Eq. (8), we have

$$\|\nabla f_S(x) - \nabla f_S(y)\| = \|P^{-1}\nabla f(x) - P^{-1}\nabla f(y)\| \le \|P^{-1}A^T A\|\|x - y\|,$$

and it holds

$$L_S = \|P^{-1}A^T A\| = \|A^T S^{-1} A\|.$$

Algorithm 2. PNPD

Choose $u_0 \in \mathbb{R}^d$, $P \in \mathbb{R}^{d\times d}$ s.p.d, $0 < \alpha < \frac{2}{L_S}$, $0 < \beta < \frac{2}{\|W\|^2}$, $k_{\max} \in \mathbb{N}$, $\{\gamma_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}_{\geq 0}$.
$u_{-1} = u_0$
$v_0^0 = 0$
for $n = 0, 1, \ldots$ **do**
  $\bar{u}_n = u_n + \gamma_n(u_n - u_{n-1})$
  $u_{n+\frac{1}{2}} = \bar{u}_n - \alpha P^{-1}\nabla f(\bar{u}_n)$
  for $k = 0, 1, \ldots, k_{\max} - 1$ **do**
    $u_n^k = u_{n+\frac{1}{2}} - \alpha W^T v_n^k$
    $v_n^{k+1} = \mathrm{prox}_{\beta\alpha^{-1}h^*}(v_n^k + \beta\alpha^{-1}W u_n^k)$
  end for
  $v_{n+1}^0 = v_n^{k_{\max}}$
  $u_n^{k_{\max}} = u_{n+\frac{1}{2}} - \alpha W^T v_n^{k_{\max}}$
  $u_{n+1} = \frac{1}{k_{\max}} \sum_{j=1}^{k_{\max}} u_n^j$
end for

3 Unrolling Architecture

In this section, we describe the proposed unrolling architecture for the PNPD algorithm and the loss function used for training. The main idea of unrolling is to truncate the iterations of an algorithm to a fixed number N. Then, the resulting truncated algorithm is interpreted as a neural network, where the parameters of the algorithm are learned through supervised training (see [20]). A further step is to not just set the parameters of the algorithm as learnable, but to obtain them as the output of a predictor network, which takes as input some information on the current state of the algorithm, such as the current iterate and the observed data. The resulting method will therefore estimate the parameters adaptively, based on the given information, being able, for example, to infer different parameters for different noise levels or different types of images. This is, for example, the approach adopted in [6] for a proximal interior point algorithm, and is the same that we adopt in this work. In particular, let us consider a single iteration of PNPD (Algorithm 2) for which the preconditioner P is defined as in Eq. (10). In this case, one step of the algorithm depends on the choice of the parameters $\lambda, \gamma, \alpha, \beta$, and ν, which are respectively the regularization, the extrapolation, the step size, the dual step size, and the preconditioner parameters. For simplicity, we will not consider the extrapolation step, i.e., we will set $\gamma = 0$. For each parameter $\theta \in \{\lambda, \alpha, \beta, \nu\}$ and for each iteration $i = 0, \ldots, N-1$, we have a convolutional predictor network

$$\mathcal{N}_{\theta_i}(b^\delta, u_i) = \theta_i,$$

which takes as input the observed data b^δ and the current iterate u_i, and outputs the parameter θ_i to be used in the i-th iteration of the algorithm. These predictor networks are then composed with the iterations of the PNPD algorithm to obtain a single iteration of the unrolling architecture, which we denote by $\mathcal{M}_i$ and which

is depicted in Fig. 1. Finally, by composing N times the base network $\mathcal{M}_i$, we obtain the global unrolling architecture, which is depicted in Fig. 2.

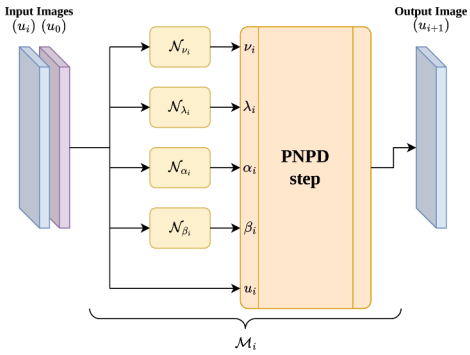


Fig. 1. Single iteration of the unrolling architecture $\mathcal{M}_i$ composed by the predictors for each parameter $\mathcal{N}_{\nu_i}, \mathcal{N}_{\lambda_i}, \mathcal{N}_{\alpha_i}, \mathcal{N}_{\beta_i}$ and by one step of the PNPD algorithm.

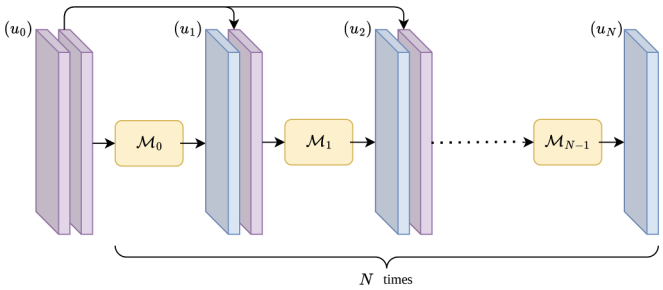


Fig. 2. Global unrolling architecture obtained by composing the base network $\mathcal{M}_i$ N times.

If the proposed architecture is used as described above for inference, there is no natural way to extend the number of iterations beyond N. For this reason, we set

$$\mathcal{N}_{\theta_i} = \mathcal{N}_{\theta_0}, \quad \forall i = 1, \ldots, N-1.$$

Thus, we force the predictor networks to learn a mapping that is independent of the iteration, based only on the current iterate and on the observed data, which allows us to naturally extend the number of iterations at inference time.

Since each $\mathcal{M}_i$ represents one iteration of the PNPD algorithm, all the intermediate iterates $u_0, u_1, \ldots, u_N$ are available as output of the unrolling architecture. This can be leveraged in the design of the loss function used for training. In

particular, the loss function is defined as a weighted average of the loss computed on each intermediate iterate, i.e.,

$$\mathcal{L}(u_0, \dots, u_{N-1}, u) = \frac{\sum_{i=0}^{N-1} w_i \, \ell(u_i, u)}{\sum_{i=0}^{N-1} w_i}, \tag{11}$$

where w_i is the weight associated with the i-th iterate, ℓ is a loss function, such as the mean squared error or the structural similarity index measure (SSIM), and u is the ground truth signal. This way of defining the loss is different from the one typically used in unrolling, where the loss is computed only on the final iterate (see, for example, [6] and [17]). The proposed approach aims to reduce the vanishing gradient problem (see [19]), which occurs when training neural networks with a large number of layers via backpropagation. By computing the loss on each iterate, this problem is attenuated, since the gradient can be back-propagated from intermediate iterates, and not only from the final one. Furthermore, this way of defining the loss stabilizes the training using information that would otherwise be ignored. By selecting appropriate weights w_i, it is possible to balance between prioritizing the final reconstruction and achieving good reconstruction quality at earlier iterations. This makes the training landscape more favorable for optimization and makes it easier to avoid local minima [11].

4 Numerical Results

In this section, we test the proposed unrolling architecture (see Sect. 3) on the image deblurring problem with TV regularization, which corresponds to the optimization problem

$$\operatorname*{argmin}_{u \in \mathbb{R}^d} \frac{1}{2} \|Au - b^\delta\|^2 + \lambda \, \mathrm{TV}(u),$$

where $A \in \mathbb{R}^{d \times d}$ is a blur operator with periodic boundary conditions, $b^\delta \in \mathbb{R}^d$ is the observed blurred and noisy image, $\mathrm{TV}(u)$ is the total variation defined as

$$\mathrm{TV}(u) = \sum_{i=1}^{d} \|(\nabla u)_i\|_2,$$

and $\lambda > 0$ is a regularization parameter.

The algorithm and the tests are implemented in Python 3.14 using PyTorch 2.10. The code is available at a public git repository[1]. The dataset used for training and testing is the BSD500 dataset [3], which consists of 500 natural images subdivided into a training set of 200 images, a validation set of 100 images, and a test set of 200 images. The observed images b^δ are generated by firstly taking a random crop of size 256×256 from an image of the dataset,

[1] https://github.com/Giuseppe499/PNPD_unrolling.

then applying a Gaussian blur with standard deviation $\sigma = 2$ pixels, and finally adding white Gaussian noise such that $\|\eta_\delta\| = 0.01\|b\|$.

The predictor networks for the parameters λ, α, β, and ν are convolutional neural networks with 2 convolutional layers. Both layers have kernel size 5, the first layer has 6 filters and the second layer has 32 filters. Each convolutional layer is followed by a ReLU activation function, and a max pooling layer with kernel size 2 and stride 2. After the convolutional layers, the output is flattened and passed through a fully connected network with a hidden layer of 128 neurons and a ReLU activation function, and an output layer with a single neuron which outputs the parameter to be used in the i-th iteration of the algorithm.

The loss function is defined as in Eq. (11) with

$$\ell(u_i, u) = -\,\mathrm{SSIM}(u_i, u),$$

where SSIM is the structural similarity index measure, and with the weights w_i defined as

$$w_i = \frac{1}{N - i},$$

thus the last iterations are more relevant than the earlier ones. The number of steps of the unrolling architecture was set to $N = 18$ for training. The training was performed for 50 epochs using the Adam optimizer with a learning rate of 10^{-4} and a batch size of 8. The best performing model on the validation set was selected for testing.

Figure 3 shows the reconstruction of some elements of the test set after 100 iterations of the proposed method. We note that, as described in Sect. 3, the number of iterations can be naturally extended beyond N at inference time, since the predictor networks are shared across iterations. It is also possible to observe that the SSIM metric for the reconstruction is higher than that of the observed image for all the shown examples, without the need for manual tuning of the parameters of the algorithm.

Figure 4 shows the evolution of the parameters and of the SSIM at inference time as a function of the iterations for the same examples shown in Fig. 3. We observe that for all the learned parameters and for a fixed image, the predictor networks tend to output almost constant values across iterations. By varying the image, we observe that the predictor networks tend to always use a step size around $\alpha = 1.5$, a dual step size around $\beta = 0.0625$, and a regularization parameter λ of the order of 10^{-3}. We recall that the theoretically sufficient condition for the convergence of the PNPD algorithm is that the gradient descent step size is $\alpha < 2/L_S = 2$, where L_S is the Lipschitz constant of ∇f_S, while the dual step size is $\beta < \|W\|^{-2} = \frac{1}{8}$, since W is the discrete gradient operator. The fact that the predicted regularization parameter λ is almost constant across different images of the test set is coherent with theoretical expectations. In particular, the blurring operator and the noise level are fixed for the test set, and therefore, the optimal regularization parameter is expected to be similar across all examples, slightly varying due to the different content of the images. We also observe that the predictor network for the preconditioner parameter ν outputs values

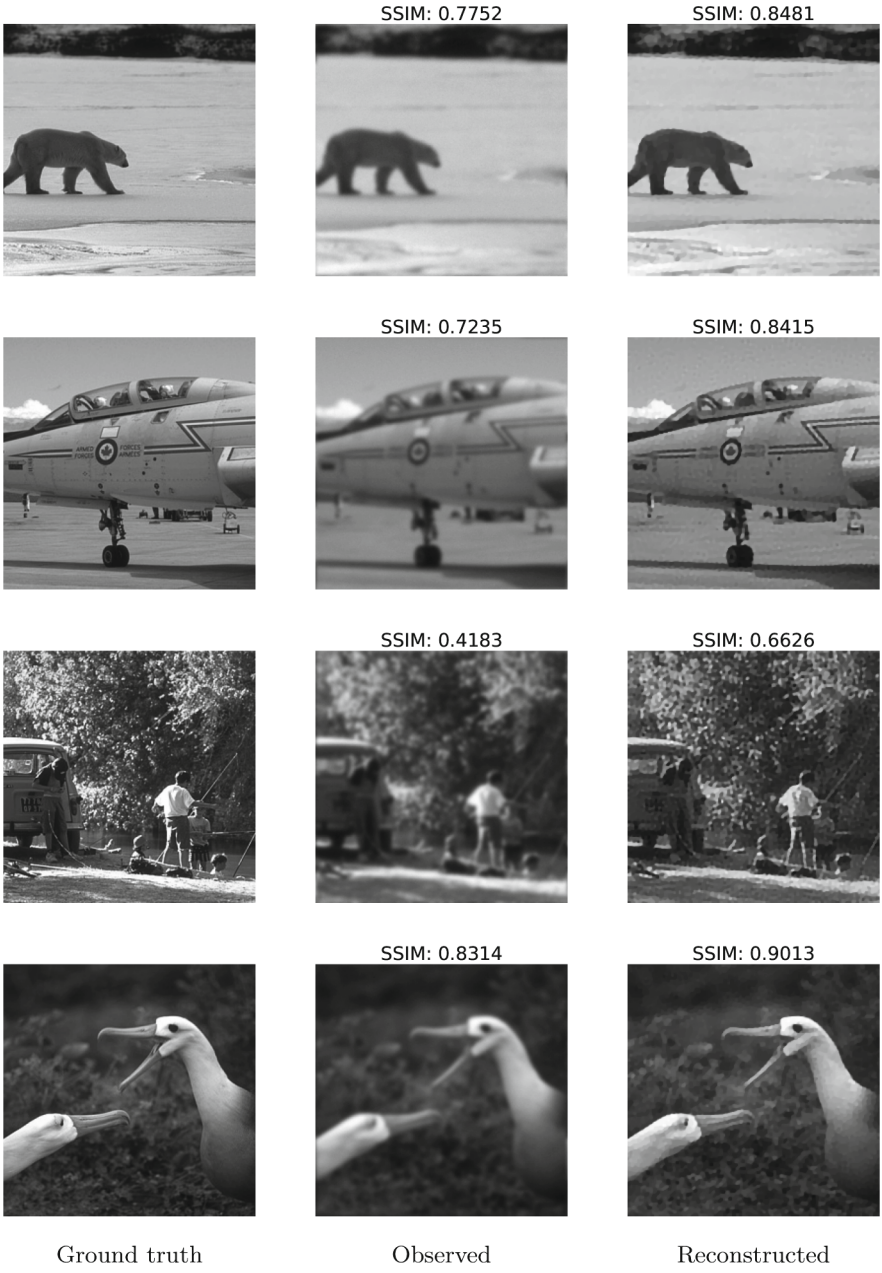


Fig. 3. Reconstruction of some images of the test set.

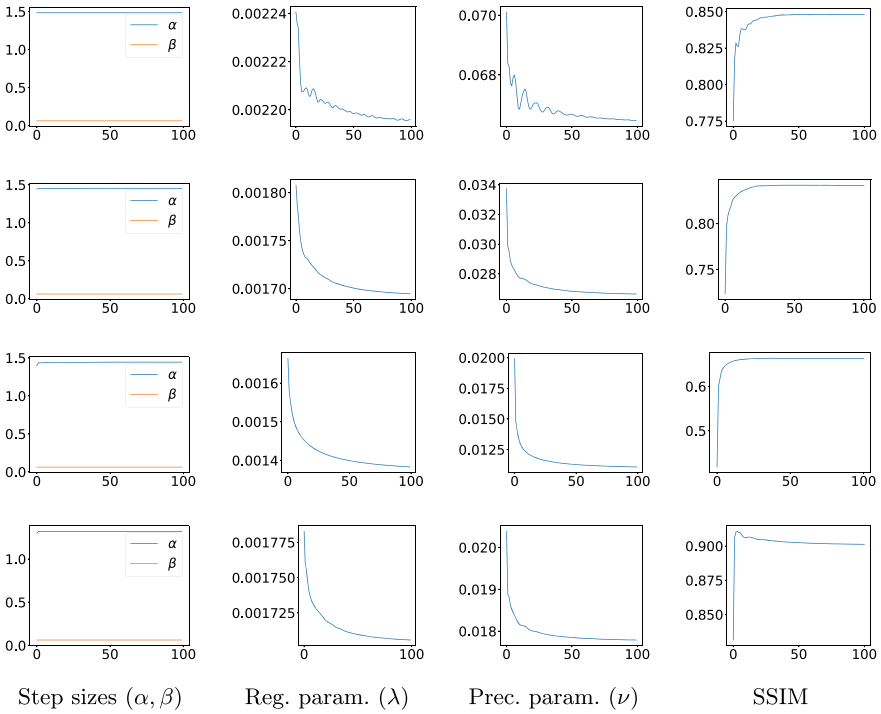


Fig. 4. Evolution of the parameters and of the SSIM as a function of the iterations for the images of the test set shown in Fig. 3.

that are in the range $[10^{-2}, 10^{-1}]$. This is consistent with the manual tuning of the parameter ν for the original PNPD algorithm, cf. [2], which led to fast convergence and good reconstruction quality for values of ν in the same range, while smaller values of ν led to instability of the algorithm, and larger values led to slower convergence, comparable to the one of the original NPD algorithm. Figure 4 also shows the evolution of the SSIM across iterations. We observe that the SSIM is monotonically increasing for the second and third images. For the first image, the SSIM exhibits slight oscillations in the initial iterations, but its overall behavior is increasing and stabilizes going on with the iterations. This may be due to ν_i being too small, which can lead to instability in the early iterations of the algorithm. For the last image, the SSIM reaches a peak at iteration 3 and then stabilizes at a slightly lower value. In this case, early stopping, for example at $i = N = 18$, could yield a slightly better reconstruction quality.

Finally, Fig. 5 shows the evolution of the negative loss function as a function of epochs and of time during training. Since the loss function is defined as the average of the negative SSIM computed on the intermediate iterates, this value can be interpreted as a measure of the reconstruction quality on the training and

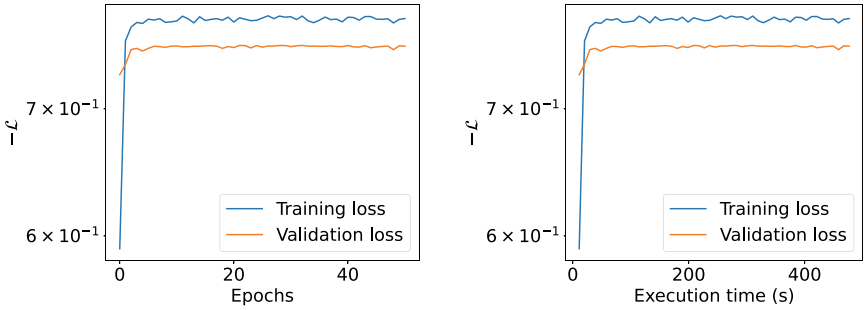


Fig. 5. Evolution of the loss function as a function of epochs (left) and of time (right).

validation sets. We observe that the negative loss increases significantly during the first epochs, while it remains nearly constant for the rest of the training.

5 Conclusion

In this work, we proposed an unrolled architecture for the PNPD algorithm, which enables the automatic estimation of the parameters involved in both the optimization problem and the algorithm, without the need for manual tuning. Numerical results on the image deblurring problem demonstrate the effectiveness of the proposed approach in terms of both reconstruction quality and parameter estimation.

Some aspects of the proposed method have not been fully explored. One of these concerns the design of the predictor networks, which in this work are simple convolutional networks, as well as the choice of input features, here limited to the observed image and the current iterate. Additional information on the state of the algorithmm, such as the dual iterates, residuals, or other features relevant for parameter estimation, could be incorporated as input to the predictor networks. Another aspect that warrants further investigation is the design of the loss function, which in this work is defined as a weighted average of the negative SSIM computed over the intermediate iterates. Different weighting strategies could be analyzed to see how they affect the training and the reconstruction quality. Finally, the training procedure of the proposed architecture could be further analyzed to improve convergence and stability, ultimately leading to enhanced reconstruction quality and more accurate parameter estimation.

Acknowledgments. The authors are partially supported by the Italian Ministry of the University and Research, through the PRIN 2022 project "Inverse Problems in the Imaging Sciences (IPIS)" (No. 2022ANC8HL), and GNCS-INdAM.

References

1. Aleotti, S., Bonettini, S., Donatelli, M., Prato, M., Rebegoldi, S.: A nested primal-dual iterated Tikhonov method for regularized convex optimization. Comput. Optim. Appl. **91**(2), 357–395 (2025). https://doi.org/10.1007/s10589-024-00613-4
2. Aleotti, S., Donatelli, M., Krause, R., Scarlato, G.: A preconditioned version of a nested primal-dual algorithm for image deblurring. J. Sci. Comput. **103**(3), 85 (2025). https://doi.org/10.1007/s10915-025-02863-8
3. Arbelaez, P., Maire, M., Fowlkes, C., Malik, J.: Contour detection and hierarchical image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **33**(5), 898–916 (2011). https://doi.org/10.1109/TPAMI.2010.161
4. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging Sci. **2**, 183–202 (2009)
5. Bertero, M., Boccacci, P.: Introduction to inverse problems in imaging. Institute of Physics Publishing, Bristol (1998)
6. Bertocchi, C., Chouzenoux, E., Corbineau, M.C., Pesquet, J.C., Prato, M.: Deep unfolding of a proximal interior point method for image restoration. Inverse Prob. **36**(3), 034005 (2020). https://doi.org/10.1088/1361-6420/ab460a
7. Bonettini, S., Loris, I., Porta, F., Prato, M., Rebegoldi, S.: On the convergence of a linesearch based proximal-gradient method for nonconvex optimization. Inverse Prob. **33**(5), 055005 (2017). https://doi.org/10.1088/1361-6420/aa5bfd
8. Bonettini, S., Prato, M., Rebegoldi, S.: A nested primal-dual FISTA-like scheme for composite convex optimization problems. Comput. Optim. Appl. **84**, 85–123 (2023)
9. Chambolle, A., Pock, T.: An introduction to continuous optimization for imaging. Acta Numerica **25**, 161–319 (2016)
10. Chen, J., Loris, I.: On starting and stopping criteria for nested primal-dual iterations. Numer. Algorithms **82**, 605–621 (2019)
11. Chen, Y., Chen, X., Maleki, A., Jalali, S.: Comprehensive examination of unrolled networks for solving linear inverse problems. Entropy **27**(9) (2025). https://doi.org/10.3390/e27090929
12. Chouzenoux, E., Pesquet, J.-C., Repetti, A.: Variable metric forward–backward algorithm for minimizing the sum of a differentiable function and a convex function. J. Optim. Theor. Appl. **162**(1), 107–132 (2013). https://doi.org/10.1007/s10957-013-0465-7
13. Combettes, P.L., Pesquet, J.C.: Proximal splitting methods in signal processing. In: Bauschke, H., Burachik, R., Combettes, P., Elser, V., Luke, D., Wolkowicz, H. (eds.) Fixed-Point Algorithms for Inverse Problems in Science and Engineering. SOIA, vol. 49, pp. 185–212. Springer, New York (2011). https://doi.org/10.1007/978-1-4419-9569-8_10
14. Daubechies, I., Defrise, M., Mol, C.D.: An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Commun. Pure Appl. Math. **57**(11), 1413–1457 (2004)
15. Engl, H.W., Hanke, M., Neubauer, G.: Regularization of Inverse Problems, 1996th edn. Mathematics and Its Applications. Springer, Dordrecht (1996). ISBN: 0-7923-4157-0
16. Golub, G.H., Heath, M., Wahba, G.: Generalized cross-validation as a method for choosing a good ridge parameter. Technometrics **21**(2), 215 (1979)
17. Gregor, K., LeCun, Y.: Learning fast approximations of sparse coding. In: Proceedings of the 27th International Conference on Machine Learning, ICML'10, pp. 399–406. Omnipress, Madison, WI, USA (2010)

18. Hansen, P.C., O'Leary, D.P.: The use of the L-curve in the regularization of discrete ill-posed problems. SIAM J. Sci. Comput. **14**(6), 1487–1503 (1993). https://doi.org/10.1137/0914086
19. Kolen, J.F., Kremer, S.C.: Gradient flow in recurrent nets: the difficulty of learning longterm dependencies, pp. 237–243 (2001). https://doi.org/10.1109/9780470544037.ch14
20. Monga, V., Li, Y., Eldar, Y.C.: Algorithm unrolling: interpretable, efficient deep learning for signal and image processing. IEEE Sig. Process. Mag. **38**(2), 18–44 (2021). https://doi.org/10.1109/MSP.2020.3016905
21. Ochs, P., Chen, Y., Brox, T., Pock, T.: iPiano: inertial proximal algorithm for nonconvex optimization. SIAM J. Imag. Sci. **7**(2), 1388–1419 (2014). https://doi.org/10.1137/130942954
22. Schmidt, M., Roux, N., Bach, F.: Convergence rates of inexact proximal-gradient methods for convex optimization. In: Shawe-Taylor, J., Zemel, R., Bartlett, P., Pereira, F., Weinberger, K. (eds.) Advances in Neural Information Processing Systems, vol. 24. Curran Associates, Inc. (2011)
23. Tikhonov, A.N., Goncharsky, A., Stepanov, V.V., Yagola, A.G.: Numerical Methods for the Solution of Ill-Posed Problems, 1995th edn. Mathematics and Its Applications. Springer, Dordrecht, The Netherlands (1995)

Numerical Aspects in Machine Learning For Dynamical Systems (NAMLYS 2026)

Intrinsic and Post-hoc Interpretability of Gaussian-Based Kolmogorov–Arnold Networks

Pasquale De Luca(✉), Emanuel Di Nardo, Emilio Garzia, Livia Marcellino, and Angelo Ciaramella

Department of Science and Technology, Parthenope University of Naples, Centro Direzionale C4, 80143 Naples, Italy
{pasquale.deluca,emanuel.dinardo,livia.marcellino, angelo.ciaramella}@uniparthenope.it, emilio.garzia001@studenti.uniparthenope.it

Abstract. Standard feedforward nets hang their nonlinearity on the nodes and keep the edges linear. Kolmogorov–Arnold Networks swap the two: each edge is now a learnable one-variable map, each node a plain summation. A small architectural move, but with a payoff we want to test here, namely, that the trained network becomes partly readable, edge by edge, without extra tooling. Our test bed is the shared-basis Gaussian variant of the KAN, recently published, and the task is regression on California Housing. We train a three-layer network with five Gaussian kernels per layer and then look at it from four angles. The first angle is the tensor of first-layer coefficients, summarized into a per-feature importance number. The second is KernelSHAP run globally; the third is LIME, pointwise; the fourth, and the one that only makes sense on a KAN, is the direct drawing of the learned edge activations as curves. Intrinsic and SHAP rankings land on the same two features at the top, MedInc and Latitude, and the same feature at the bottom (Population). They shuffle the middle, and that shuffle is diagnostic: coefficient size reports on encoding strength in layer zero, SHAP on end-to-end marginal effect. The plotted edge activations confirm what domain knowledge would suggest: income acts roughly monotonically on price, geography acts locally, crowding has a threshold effect. Taken together, the four views make the model easier to audit than a comparable black box would be.

Keywords: Kolmogorov–Arnold networks · explainable AI · radial basis functions · SHAP · feature importance · tabular regression

1 Introduction

A wrong prediction in loan approval, in property valuation, in medical triage or cancer prognosis [17] is not just a misclassification, it has downstream costs someone has to pay. In those settings, being able to audit *why* the model predicted

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 173–183, 2027.
https://doi.org/10.1007/978-3-032-30530-5_11

what it did weighs about as much as the prediction's accuracy on a held-out fold. The trouble is that modern deep networks, stacking nonlinearity after nonlinearity, are hard to audit in any direct way; so the machine-learning community has taken two routes to address the problem. One route is external: keep the black box, and build tools around it that probe and summarize its behavior, locally or globally [1,2]. The other is internal: rethink the architecture so the model is readable from the outset [3].

Kolmogorov–Arnold Networks [4] belong to the second tradition, and they take their justification from a theorem by Kolmogorov [5]. Any continuous $f\colon [0,1]^d \to \mathbb{R}$, the theorem says, factors through univariate maps:

$$f(x_1,\ldots,x_d) = \sum_{q=0}^{2d} \Phi_q\Big(\sum_{p=1}^{d} \phi_{q,p}(x_p)\Big), \tag{1}$$

where every $\phi_{q,p}$ and every Φ_q is a scalar-to-scalar function. A KAN turns (1) into an architecture: edges carry the univariate functions, and those functions are learned as coefficients in some basis. For us, the relevant point is that each learned $\phi_{q,p}$ takes a single real number and returns a single real number. Which means it can be drawn. And drawn things, unlike opaque tensors, admit inspection.

The B-spline parameterization of the original KAN was recently replaced in [6], with Gaussian radial basis functions under a sharing rule: all edges of a layer pull from the same small pool of kernels. The benefit is a drop in parameter count from $\mathcal{O}(3KW)$ down to $\mathcal{O}(KW + 2LK)$, no loss of Sobolev convergence order, and clear gains on physics-informed benchmarks [18], especially hyperbolic PDEs or complex reaction-diffusion systems [15,16], where the C^∞ smoothness of Gaussians matters for high-order derivatives. Interpretability, though it was implicit in the design, was not the focus of that paper.

1.1 Contribution and Outline

This paper takes up interpretability explicitly. We pick California Housing because $d = 8$ is small, the features are interpretable to a human reader (income, latitude, age of the housing stock), and the target is familiar. Our protocol combines the two XAI traditions without privileging either: we read the trained coefficients where they talk, we probe the model with SHAP and LIME where they do, and we plot the edge activations because the architecture lets us. The central question is whether the four viewpoints, put side by side, tell the same story, or a richer one when they disagree.

Section 2 fixes the notation, recalls the shared-basis construction, and isolates the two interpretability handles its structure gives us for free. Section 3 describes the protocol and the four-level pipeline. Section 4 walks through the results, with attention to the places where the four levels do not line up. Section 5 summarizes, and flags a few open ends.

2 Mathematical Background

The purpose of this section is twofold: to state what a shared-basis RBF-KAN is, exactly, and to identify two observable quantities its structure gives us at no extra cost.

2.1 Shared-Basis RBF-KAN Architecture

Every layer $l \in \{0, \dots, L-1\}$ owns a pool of K Gaussian kernels, parametrized by centres $\boldsymbol{\mu}^{(l)} = (\mu_1^{(l)}, \dots, \mu_K^{(l)})$ and widths $\boldsymbol{\sigma}^{(l)} = (\sigma_1^{(l)}, \dots, \sigma_K^{(l)})$. To stop the widths from collapsing to zero, which would blow up the exponent in (2), they are clamped at $\sigma_{\min} = 10^{-6}$ during training [6]. The edge from input neuron j to output neuron i at layer l carries the activation

$$\phi_{ij}^{(l)}(x) = \sum_{k=1}^{K} c_{ijk}^{(l)}\, \psi_k^{(l)}(x), \qquad \psi_k^{(l)}(x) = \exp\left(-\frac{(x-\mu_k^{(l)})^2}{2(\sigma_k^{(l)})^2}\right), \tag{2}$$

with learnable scalar coefficients $c_{ijk}^{(l)} \in \mathbb{R}$ and shared kernels $\psi_k^{(l)}$. Propagation at layer l sends $\mathbf{h}^{(l)} \in \mathbb{R}^{d_l}$ to $\mathbf{h}^{(l+1)} \in \mathbb{R}^{d_{l+1}}$ through

$$h_i^{(l+1)} = \sum_{j=1}^{d_l} \phi_{ij}^{(l)}\big(h_j^{(l)}\big), \qquad i = 1, \dots, d_{l+1}. \tag{3}$$

The sharing is what makes this parsimonious. All $d_l \cdot d_{l+1}$ edges of a layer plug into the same K Gaussians; per-edge variation lives entirely in the coefficient vector $\mathbf{c}_{ij}^{(l)} \in \mathbb{R}^K$. Summing up: $\sum_{l=0}^{L-1}(K \cdot d_l \cdot d_{l+1} + 2K)$ parameters, i.e. $\mathcal{O}(KW + 2LK)$ with $W = \sum_l d_l d_{l+1}$.

2.2 Intrinsic Interpretability Mechanisms

Two quantities drop out of the formulae above with no added machinery.

Coefficient-Based Feature Importance. The first-layer tensor $\mathbf{C}^{(0)} \in \mathbb{R}^{d_0 \times d_1 \times K}$ has, at entry (i, j, k), the weight that input feature j puts on hidden unit i through kernel k. A simple per-feature summary is the mean absolute value,

$$\mathcal{I}_j^{\mathrm{RBF}} = \frac{1}{d_1 K} \sum_{i=1}^{d_1} \sum_{k=1}^{K} |c_{ijk}^{(0)}|, \tag{4}$$

which tells us, roughly, how loud the first-layer writing of feature j into the hidden state is. No sampling, no surrogate, no perturbation: one reads (4) off the trained weights.

Activation Function Visualization. Each $\phi_{ij}^{(l)}$ is a scalar function of a scalar variable, written in a known Gaussian basis with learned coefficients, so plotting $x \mapsto \phi_{ij}^{(l)}(x)$ is not a proxy for an explanation, it is the actual trained object. Drawing all edges of a three-layer network would not be readable, though, so we rank them by their ℓ^2 coefficient norm

$$\nu_{ij}^{(l)} = \Big(\sum_{k=1}^{K} (c_{ijk}^{(l)})^2\Big)^{1/2} \tag{5}$$

and keep the top few per layer. The next bound says the ranking is not arbitrary.

Proposition 1 (Pointwise activation bound). *If* $\|\psi_k^{(l)}\|_\infty = 1$ *for every* k, l, *then for any edge* (i, j) *in layer* l

$$\sup_{x \in \mathbb{R}} |\phi_{ij}^{(l)}(x)| \;\le\; \sqrt{K}\, \nu_{ij}^{(l)}. \tag{6}$$

Proof. Apply Cauchy–Schwarz to the sum in (2): $|\phi_{ij}^{(l)}(x)| \le \nu_{ij}^{(l)} \big(\sum_k |\psi_k^{(l)}(x)|^2\big)^{1/2}$. Since $|\psi_k^{(l)}(x)| \le 1$ everywhere, the inner sum is at most K and its square root is at most $\sqrt{K}$.

So $\nu_{ij}^{(l)}$ sets, up to a $\sqrt{K}$ factor, a ceiling on the loudness of edge (i, j). Ranking by $\nu_{ij}^{(l)}$ and keeping the top edges keeps, by construction, the edges that have room to matter most.

3 Explainability Framework

Section 2 gave us the parts. This section describes the pipeline that puts them to work. The design principle is a simple one: no single explainer is trusted on its own; four structurally different tools are run on the same trained model, and we look at where they converge and where they do not.

3.1 Experimental Setup

California Housing [7] is the data source, $n = 20{,}640$ rows, $d = 8$ numeric features (MedInc, HouseAge, AveRooms, AveBedrms, Population, AveOccup, Latitude, Longitude), and a target representing median house value in units of \$100k.

The model is a three-hidden-layer RBF-KAN of widths $[8, 8, 8, 1]$, $K = 5$ Gaussian kernels per layer, `kan_layer = rbf_linear`, initialization noise $\varepsilon_{\text{noise}} = 0.1$. Fifty L-BFGS epochs, learning rate $\eta = 1.0$, batch size $B = 512$, MSE loss. Inputs and target are standardized. Split 80/20, seed fixed.

3.2 Four-Level Interpretability Pipeline

From most structural to least, the four levels are these.

Level 1: Intrinsic RBF Coefficient Importance. Compute $\mathcal{I}_j^{\mathrm{RBF}}$ from (4) directly on the trained first-layer tensor. Nothing else to run.

Level 2: Global Post-Hoc Attribution via KernelSHAP. KernelSHAP [1] approximates Shapley values [8] by weighted least squares over sampled coalitions. We take $N_{\mathrm{bg}} = 100$ background points, explain $N_{\mathrm{test}} = 300$ test points, and use $n_{\mathrm{samples}} = 128$ kernel evaluations per explained instance. Mean absolute SHAP values $\overline{|\phi_j|}$ give the global ranking.

Level 3: Local Post-hoc Attribution via LIME. LIME [2] fits a sparse linear model locally around the point being explained. We pick three representative test points, the lowest, the median, and the highest predicted values, and explain each with 1000 perturbations.

Level 4: Structural Visualization. Two plots: first, $\phi_{ij}^{(l)}$ for the top-4 edges of each layer (by the norm in (5)), shown together with their individual weighted basis components; second, a per-layer heatmap of $\frac{1}{d_{l+1}} \sum_i |c_{ijk}^{(l)}|$ over input neuron j and knot index k. This is the level that requires the KAN formulation; nothing here has a direct counterpart for an MLP.

4 Results

We cover two things in this section: how the network fits (briefly, to establish that we have a sensible model to explain), and what the four interpretability levels say about it (at length, since that is the contribution).

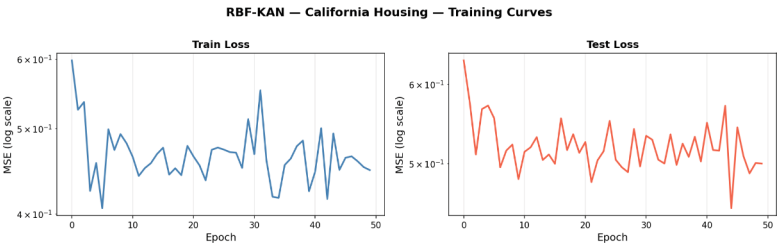


Fig. 1. Training and test MSE (log scale), 50 L-BFGS epochs. Jaggedness is a property of the optimiser, not of the loss surface.

4.1 Predictive Performance and Training Dynamics

After fifty L-BFGS epochs the test MSE lands near 4.7×10^{-1} in standardized units, enough for the model to have learned something non-trivial about the data, not so much that it has memorized it. The loss curves in Fig. 1 are jagged rather than smooth. This is expected: L-BFGS relies on a line search that occasionally overshoots before correcting, which is the cost one pays for using curvature information. Train and test curves move together throughout; the model does not overfit. Total parameters: about 1,070, noticeably fewer than what a dense MLP for this task would need, which is one of the things the sharing rule was designed to deliver [6].

4.2 Feature Importance: Intrinsic vs Post-Hoc Rankings

Table 1 lines up the four rankings. Intrinsic score first: MedInc leads at $\mathcal{I}^{\mathrm{RBF}} = 0.1911$, Latitude follows at 0.1656, AveOccup third at 0.1567. SHAP reorders: Latitude takes first place at $\overline{|\phi|} = 0.7264$, Longitude second at 0.5886, MedInc drops to third at 0.3170.

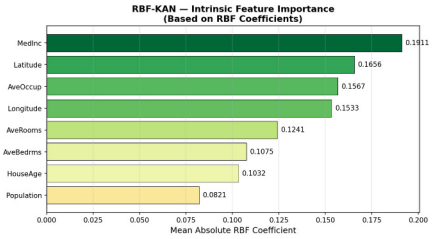
Table 1. Feature importance rankings across the four interpretability levels. Rank 1 is the most influential. The LIME column shows the dominant rank over three representative instances (low/median/high prediction).

Feature	$\mathcal{I}^{\mathrm{RBF}}$ Rank	SHAP Rank	LIME (dominant)
MedInc	1	3	2–3
Latitude	2	1	2
AveOccup	3	4	1
Longitude	4	2	3
AveRooms	5	5	4–5
AveBedrms	6	7	5–6
HouseAge	7	6	6–7
Population	8	8	7–8

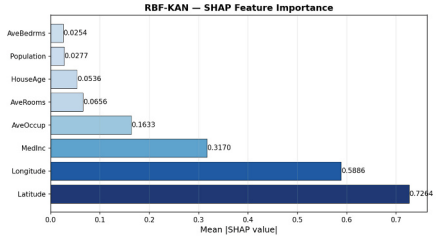
The two rankings are measuring different objects, and the difference is informative. $\mathcal{I}_j^{\mathrm{RBF}}$ weighs coefficients in the first layer: it records how emphatically feature j is injected into the initial representation. SHAP measures something else: once all three layers have run, how much does feature j shift the output at the margin. Latitude and Longitude carry only moderate first-layer weight, but the composition of the subsequent layers amplifies their effect, which is exactly what a high SHAP value records. MedInc, conversely, enters with large first-layer coefficients but part of its weight ends up baked into interactions rather than into marginal contribution, so SHAP sees less of it at the output end.

Despite the reshuffle in the middle of the ranking, the two views agree on the endpoints: MedInc and Latitude are in the top three for both, and Population

sits at the bottom for both. Where they diverge, on Longitude, AveOccup, the internal ordering of the geographic variables, the divergence is not noise, it is the signature of multi-layer interaction one would miss looking at any single view alone (Fig. 2).



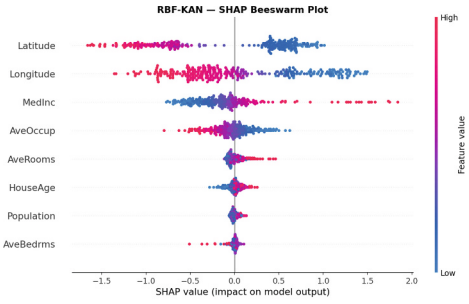
(a) Intrinsic importance (mean $|c_{ijk}^{(0)}|$).



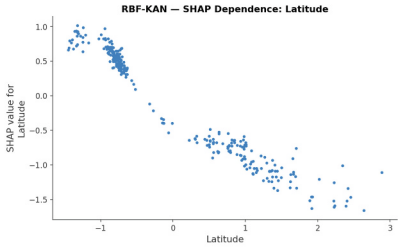
(b) SHAP global importance (mean $|\phi_j|$).

Fig. 2. Intrinsic coefficient importance (a) versus KernelSHAP (b). Both put MedInc and Latitude in the top tier. They disagree on where the two geographic variables belong.

SHAP also has a sign, which the bar plot throws away. The beeswarm in Fig. 3a brings it back. High Latitude values, Northern California, pull the prediction down; high MedInc values push it up. The Latitude dependence plot (Fig. 3b) shows the shape of that pull: smooth, monotonic, nearly linear over the observed range. The network, then, has not learned a threshold or a partition of the state into regions, but a continuous geographic gradient. None of this should surprise anyone who has looked at California house prices recently; that non-surprise is the point. A model whose explanations violate common knowledge about its domain would be a model one should hesitate to deploy.



(a) SHAP beeswarm plot.



(b) Dependence plot for Latitude.

Fig. 3. Directional SHAP views. (a) Beeswarm, showing monotonic trends for Latitude and MedInc. (b) Latitude dependence plot, a smooth near-linear descent, matching the known north–south gradient in California prices.

At the local level, LIME confirms the global picture one instance at a time (Fig. 4). AveOccup and the two geographic coordinates dominate the explanation for the low-value, median-value, and high-value test points alike; they do not drop in and out across the prediction range the way a less robust importance signal would. MedInc flips sign between the low and high instances, negative in the former, positive in the latter, which is what one expects of a feature that drives the outcome monotonically: where the prediction is low, the feature pushed it low; where it is high, the feature pushed it high. The sign flip at the local level is consistent with the monotonicity SHAP picked up globally, which is the sort of internal consistency one wants from a model one is about to trust.

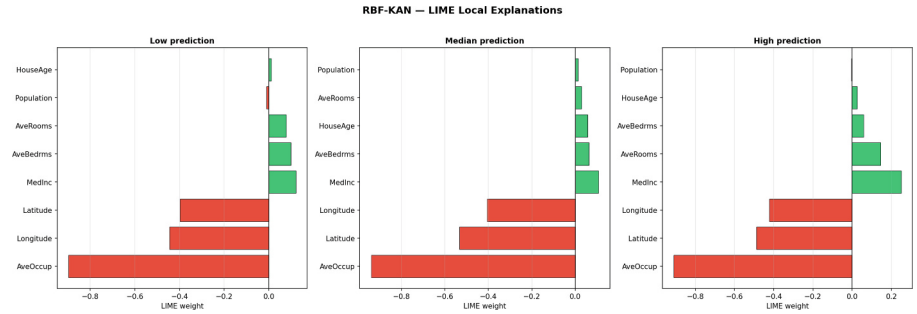


Fig. 4. LIME local explanations, three instances (low, median, high prediction). AveOccup and the geographic variables dominate in all three. MedInc flips sign cleanly between the low and high cases.

Level 4 is where the KAN formulation really earns its keep. Figure 5 plots the top four edges per layer as honest-to-goodness univariate curves, each the superposition of the five weighted Gaussians the edge has learned. First-layer Latitude→out is a single clean Gaussian peak, geographically, the Bay Area bump sitting in the region of the feature axis where the most expensive submarket lives once features have been standardized. MedInc→out is broader and roughly monotonic, which is the shape one wants from a feature that tracks the target linearly. AveOccup→out is a sharp peak near the standardized mean; past the peak the curve descends sharply, which reads cleanly as a threshold effect: once occupancy exceeds a certain level, it starts to pull prices down.

The second and third rows of Fig. 5 are busier. The shapes there are multimodal, with kinks and sharp transitions. That is the composition step doing its work: subsequent layers take the features-wise transforms of the first layer and combine them into higher-order structure. What matters for interpretability is that this structure is still, at every depth, a sum of univariate transforms of scalar arguments, so the decomposition of (1) holds through the entire network, not just in the first layer.

The heatmaps in Fig. 6 add a complementary picture. In layer 0, knot 0 has large coefficients for MedInc and Latitude; knot 2 is where AveOccup and

AveRooms put most of their weight. Different knots serve different features, and not uniformly. This specialization is possible precisely because the kernels are shared: all edges pull from the same pool of five Gaussians, so the coefficient pattern tells us which pool member each feature calls on, and how strongly.

5 Conclusions

Four windows on the same trained network: coefficients, SHAP, LIME, activation plots. None of them is redundant with the others. Coefficients read the encoding

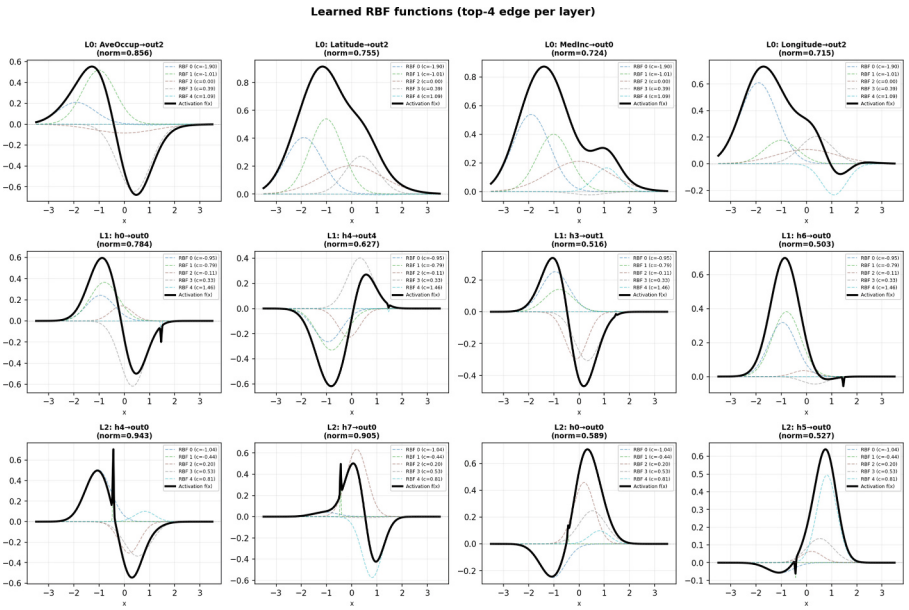


Fig. 5. Learned RBF activation functions, top-4 edges per layer (by coefficient norm). Dashed: the individual weighted Gaussian components. Solid: their sum, which is the activation the network actually applies on that edge.

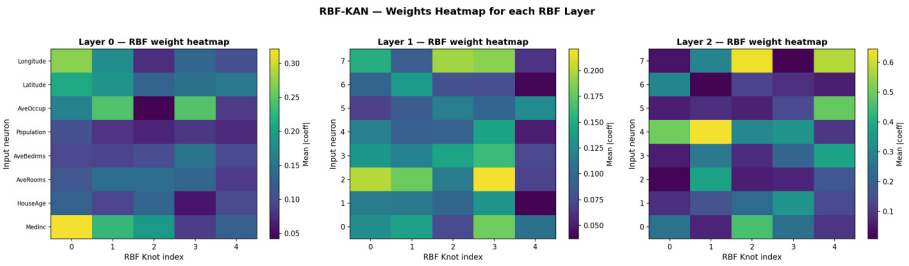


Fig. 6. Coefficient heatmaps, one per layer: mean absolute coefficient over input neurons (rows) and knot indices (columns). Unevenness is knot specialization.

stage; SHAP reads the output margin; LIME zooms in on single points; the activation plots show the nonlinearities as drawings one can inspect. When the four windows agree, the attribution is on solid ground. When they disagree, as they do on the middle of our ranking, or on the relative importance of the two geographic coordinates, the disagreement is informative, pointing to multi-layer structure that no single window would have revealed.

What sets this exercise apart from standard XAI on a black box is the last window. SHAP and LIME summarize influence as numbers, one per feature per instance; the KAN, by construction, preserves the functional form itself, and lets us look at it. That is what Rudin [3] had in mind when she argued for intrinsically interpretable models over post-hoc explainers: not more scalars to read, but the actual learned function to examine.

5.1 Limitations and Outlook

Eight features is friendly. Scaling activation plots to $d \gg 10$ is not straightforward, first-layer edges grow as $d \times d_1$, and at some point the human reader's attention runs out even though the mathematics is unchanged. One natural response is to use the norm (5) as a pruning criterion: apply Prop. 1 to filter out edges whose ceiling is below some threshold, and plot only what remains. A second limitation is that $\mathcal{I}_j^{\mathrm{RBF}}$ reads the first layer alone; a multi-layer variant, along the lines of a product of edge norms on paths from input to output, would plausibly track SHAP more closely, and would be worth developing. Beyond tabular regression, a natural follow-up is to take this analysis back to the physics-informed setting of [6,18], particularly for modeling complex dynamical systems and PDEs [15,16], where the shapes of the learned activations could be compared not merely against domain intuition but against the differential operator the network is approximating, a more stringent interpretability test than one can set up on Housing data.

Acknowledgments. This work was partially supported by INdAM - GNCS project 2026 "Metodi Numerici e Machine Learning per Sistemi Dinamici Complessi: Identificazione, Approssimazione"', ref. no. CUP E53C25002010001. P. De Luca, E. Di Nardo, L. Marcellino and A. Ciaramella are members of INdAM-GNCS.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Lundberg, S.M., Lee, S.-I.: A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems, vol. 30 (2017)
2. Ribeiro, M.T., Singh, S., Guestrin, C.: "Why should I trust you?": explaining the predictions of any classifier. In: 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1135–1144 (2016)

3. Rudin, C.: Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Mach. Intell. **1**(5), 206–215 (2019)
4. Liu, Z., et al.: KAN: Kolmogorov–Arnold networks. arXiv preprint arXiv:2404.19756 (2024)
5. Kolmogorov, A.N.: On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. Dokl. Akad. Nauk SSSR **114**, 953–956 (1957)
6. De Luca, P., Di Nardo, E., Marcellino, L., Ciaramella, A.: Stable and efficient Gaussian-based Kolmogorov-Arnold networks. Mathematics **14**(3), 513 (2026)
7. Pace, R.K., Barry, R.: Sparse spatial autoregressions. Stat. Probabil. Lett. **33**(3), 291–297 (1997)
8. Shapley, L.S.: A value for n-person games. In: Kuhn, H.W., Tucker, A.W. (eds.) Contributions to the Theory of Games, vol. 2, pp. 307–317. Princeton University Press (1953)
9. Molnar, C.: Interpretable Machine Learning, 2nd edn. (2022). https://christophm.github.io/interpretable-ml-book/
10. Park, J., Sandberg, I.W.: Universal approximation using radial-basis-function networks. Neural Comput. **3**(2), 246–257 (1991)
11. Schaback, R.: Error estimates and condition numbers for radial basis function interpolation. Adv. Comput. Math. **3**, 251–264 (1995)
12. Wendland, H.: Scattered Data Approximation. Cambridge University Press (2005)
13. Buhmann, M.D.: Radial Basis Functions: Theory and Implementations. Cambridge University Press (2003)
14. Lipton, Z.C.: The mythos of model interpretability. Queue **16**(3), 31–57 (2018)
15. De Luca, P., Marcellino, L.: Conservation law analysis in numerical schema for a tumor angiogenesis PDE system. Mathematics **13**(1), 28 (2025). https://doi.org/10.3390/math13010028
16. Cardone, A., De Luca, P., Galletti, A., Marcellino, L.: Solving Time-Fractional reaction-diffusion systems through a tensor-based parallel algorithm. XXPhys. A **611**, 128472 (2023)
17. Ciaramella, A., Di Nardo, E., Terracciano, D., et al.: A new biomarker panel of ultraconserved long non-coding RNAs for bladder cancer prognosis by a machine learning based methodology. BMC Bioinformatics **23**(Suppl 6), 569 (2022). https://doi.org/10.1186/s12859-023-05167-6
18. Colace, F., Conte, D., Pagano, G., Paternoster, B., Valentino, C.: Physics-informed neural networks for a Lithium-ion batteries model: a case of study. Adv. Comput. Sci. Eng. (ACSE) **2**(4) (2024)

A GPU Accelerated DCT Implementation for Image Compression

Angelamaria Cardone[1(✉)] and Gerardo Di Pascale[2]

[1] Department of Mathematics, University of Salerno, Via Giovanni Paolo II n. 132, 84084 Fisciano, Italy
ancardone@unisa.it

[2] Department of Computer Science, University of Salerno, Via Giovanni Paolo II n. 132, 84084 Fisciano, Italy
https://docenti.unisa.it/angelamaria.cardone

Abstract. The Discrete Cosine Transform (DCT) is a cornerstone of the JPEG standard, nevertheless its direct implementation entails significant computational complexity for full-image processing, due to intensive matrix operations. Building upon the methodology proposed by Haweel *et al.* in 2016, which utilizes a specific transformation matrix to streamline block-based DCT through matrix multiplication, this work proposes an optimized parallel version designed for GPUs. This implementation leverages advanced strategies, including memory coalescing, the reduction of thread divergence, and the efficient management of GPU memory hierarchies. Experimental results demonstrate that the optimized algorithm achieves a remarkable speed-up of 340× compared to traditional CPU-based approaches. Furthermore, the implementation significantly outperforms some existing parallel solutions for GPU, reducing execution time by up to 31% compared to an efficient CUDA-based algorithm and by over 96% compared to the standard cuBLAS library. These performance gains are especially pronounced when processing high-resolution images, highlighting the scalability and computational efficiency of the proposed approach for large-scale visual data.

Keywords: image compression · DCT · GPU computing

1 Introduction

Image compression is one of the keys to achieve speed and efficiency of many applications in the digital world. Transmitting compressed images means "transmit less" and thus be faster. Archiving compressed images means "archive more", hence efficient use of storage. DCT is widely used in the JPEG image compression standard. However, its direct computation has high computational complexity, as it is based on multiplications between very large matrices. The computation of a two-dimensional DCT, using standard procedures, requires a computational cost of $O(N^3)$ for the entire image.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 184–198, 2027.
https://doi.org/10.1007/978-3-032-30530-5_12

In the last decades, GPU computing has emerged as a powerful and highly versatile approach for tackling numerical problems of scientific calculus, which require a huge computational effort. Within this class, there are problems involving large data sets and/or characterized by considerable computational complexity. Just to make some examples: the solution of differential problems of high dimension [3–5,8,11]; images elaboration, as image denoising [6,7], image compression [9,10,16,17] and other issues related to images [13,14].

This paper deals with the study, development and optimization of an advanced image compression algorithm based on the Discrete Cosine Transform (DCT) and their optimized parallel implementation for GPUs. The main goal is to improve computational efficiency and reduce processing time by leveraging the massive parallelism inherent in the CUDA architecture.

In the recent literature, several approaches have been proposed so far for an efficient image compression based on DCT. Most of them rely on accurate and cheap approximations of the transformation matrix [1,2,10] and also suitable for parallel implementations [9]. Following the methodology introduced by Haweel *et al.* in [9], the image is divided into blocks of fixed size and then the approximated DCT of each block is computed, which requires matrix multiplications involving the image blocks, the approximated transformation matrix and its transpose. The considered algorithm has been parallelized and optimized for the implementation on the GPU architecture.

The article is organized as follows. Section 2 illustrates the image compression method based on DCT, and briefly describes some variants of this methods. In Sect. 3 we explain the GPU implementation we propose, with the optimization strategies we adopt to increase efficiency. Section 4 includes some numerical experiments carried out on significant test examples from the specialized literature and measure the accuracy and efficiency of our algorithm, also by comparing it with other existing algorithms. Finally, some concluding remarks are reported in Sect. 5.

2 Image Compression by DCT

Established in 1992, JPEG was the first international image compression standard based on the Discrete Cosine Transform (DCT) [12]. The complete encoding and decoding process involves several stages, as illustrated in Fig. 1.

JPEG compression operates on 8×8 pixel blocks, applying a sequence of transformations to each block independently, to achieve an efficient data representation. **The Forward DCT** (FDCT) step transforms image blocks into the frequency domain, obtaining the coefficient matrix. This transformation itself is lossless. Then, each of the 64 coefficients undergoes a uniform quantization using a carefully designed 64-element quantization matrix Q. The process consists of the integer division with rounding of the coefficients by the quantization matrix. Quantization introduces loss by discarding perceptually less significant details.

The two dimensional DCT of a matrix x of dimension $N \times N$, is the matrix

$$X = T_{DCT}^T x T_{DCT},$$

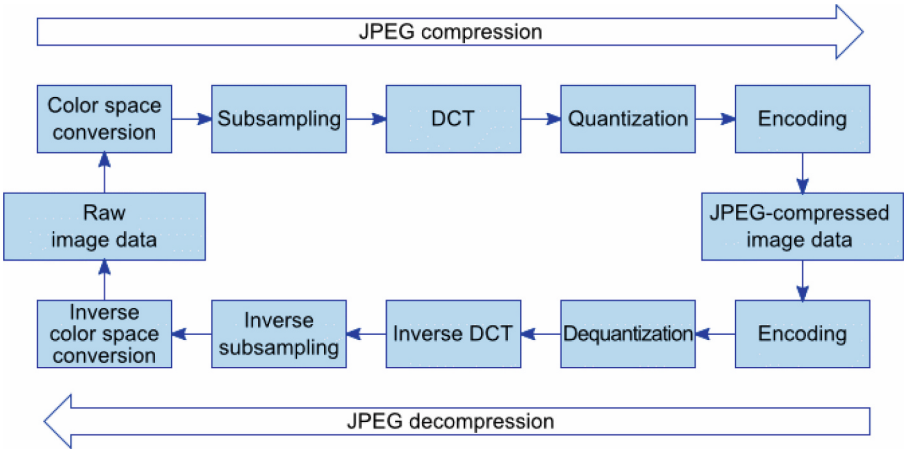


Fig. 1. JPEG Encoding Process scheme.

where A^T denotes the transpose of matrix A, and

$$T_{DCT}(u,v) = \alpha(u)\alpha(v) \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \cos\left[\frac{\pi(2i+1)u}{2N}\right] \cos\left[\frac{\pi(2j+1)v}{2N}\right],$$

$$\alpha(u) = \begin{cases} \frac{1}{\sqrt{N}}, & \text{if } u = 0, \\ \frac{2}{\sqrt{N}}, & \text{otherwise} \end{cases}$$

for $u, v = 1, \ldots, N$. We observe that the transformation matrix T_{DCT} is orthogonal, i.e. $T_{DCT}^{-1} = T_{DCT}^{T}$.

In 8-bit precision images, pixel intensities range from 0 to 255. Prior to the DCT transformation, a level shift is performed by subtracting 128 from each pixel value. This centers the dynamic range around zero, shifting it to $[-128, 127]$, which is the standard input format for DCT-based compression like JPEG.

Then, the DCT transformation is applied and a matrix X is obtained. Given a certain quantization matrix Q, each value of X is replaced by the nearest integer of the fraction

$$\frac{X_{ij}}{Q_{ij}}.$$

The inverse process is applied for the decompression of the image.

A number of studies on the DCT-based compression methods tried to identify a transformation matrix T, which accurately approximates the matrix T_{DCT} and reduces the computational effort of the overall coding and encoding process. The goal is to balance accuracy and performance, ensuring that the approximated transformation retains the essential properties of the original DCT while reducing complexity for practical applications. A fundamental property to maintain is orthogonality, which significantly streamlines the decompression process.

Under this assumption, the reconstruction $X = T^T CT$ follows directly from the forward transform $C = TXT^T$, without requiring explicit matrix inversion.

One of the first approximations was proposed by Haweel in [10], where the sign function is applied to the T_{DCT} matrix:

$$T_{SDCT}(u,v) = \frac{1}{\sqrt{N}} \operatorname{sign}\{T_{DCT}(u,v)\}$$

with the signum function operator sign defined as:

$$\operatorname{sign}\{x\} = \begin{cases} +1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases}$$

Nevertheless, this approach does not preserve orthogonality. Later, Bouguezel *et al.* in [2] improved this approximation by introducing the orthogonal matrix T shown in the next section.

3 Proposed GPU Algorithm for Image Compression by DCT

This section illustrates first the mathematical formulation of the DCT-based image compression method we consider. Then, we describe in detail the strategies we introduce for the GPU implementation, in order to improve efficiency.

3.1 Mathematical Formulation

Here, we explain the mathematical foundation underlying our algorithm, which exploit the matrix formulation of the DCT. Specifically, it is possible to express the DCT of an image block X, of size 8×8, in the form $T \cdot X \cdot T^T$, where T is a transformation matrix specially designed in [2]. This formulation facilitates an efficient implementation on GPUs, thanks to the inherent parallelism of matrix operations.

The transformation matrix used is given by:

$$T = \hat{D}\hat{T}, \tag{1}$$

with

$$\hat{T} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & -1 & -1 \\ 1 & 0.5 & -0.5 & -1 & -1 & -0.5 & 0.5 & 1 \\ 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0.5 & -1 & 1 & -0.5 & -0.5 & 1 & -1 & 0.5 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$\hat{D} = \operatorname{diag}\left(\frac{1}{2\sqrt{2}}, \frac{1}{2}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{2}}, \frac{1}{2\sqrt{2}}, \frac{1}{2}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{2}}\right)$$

We highlight that the matrix T satisfies the orthogonality property, i.e., $T^{-1} = T^T$. This property allows the same matrix to be used for both encoding and decoding. For an image block X of size 8×8, the transformation is applied as follows:

$$C = TXT^T = \hat{D}\left(\hat{T}X\hat{T}^T\right)\hat{D}^T,$$

and the reconstruction is given by:

$$X = T^T CT = \hat{T}^T\left(\hat{D}^T C\hat{D}\right)\hat{T}.$$

3.2 Algorithm Overview

The key stages of a generic compression algorithm that follow JPEG standard include:

1. **Sub-sampling:** Each element of the image matrix is subtracted by a scalar of 128.
2. **DCT Application:** The DCT is performed on each block X_b of size 8×8 of the image matrix X, obtaining the coefficients matrix C.
3. **Quantization:** The resulting coefficients are quantized using a predefined matrix Q.

Conversely, the decompression process reverses these operations:

4. **De-quantization:** The quantized coefficients are de-quantized.
5. **Inverse DCT:** The inverse transformation is applied for each block C_b of the coefficient matrix C.
6. **Up-sampling:** Pixel values are restored by adding 128 to the matrix entries.

3.3 Proposed GPU Implementation

In this section, we present the **HpApprDCT** algorithm and its implementation. Developed from scratch, this method is designed to be the fastest approach while incorporating an extended optimization phase inspired by earlier work [9].

The CUDA programming model follows a host-device architecture, where the CPU acts as the host and the GPU as the device. On the host side, data are first acquired and processed sequentially before being transferred to the device's global memory for massively parallel execution. Parallel operations on the GPU are executed through kernels, which are C functions running concurrently on multiple CUDA threads. These threads are organized into blocks, which can be organized into grids in up to three dimensions, with all threads executing the same kernel as shown in Fig. 2.

The HpApprDCT is applied to images by dividing the image into 8×8 blocks. Image block processing is performed independently, allowing all blocks to be processed efficiently in parallel and distributed across the GPU's parallel computing units.

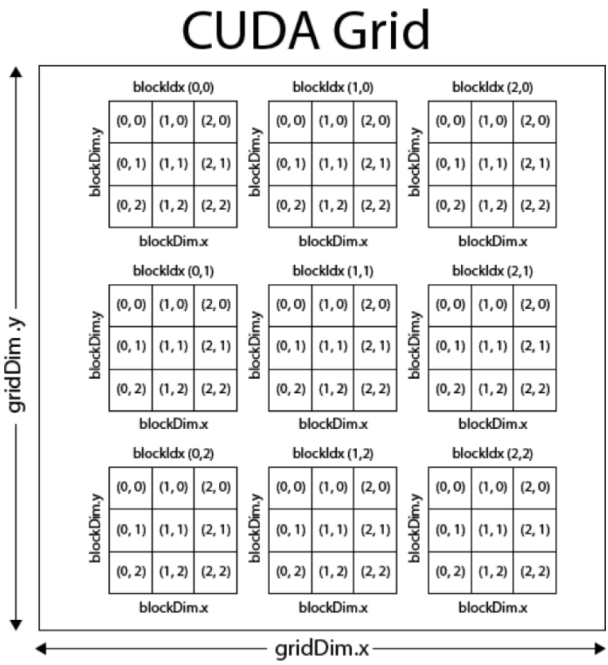


Fig. 2. Threads e Blocks within CUDA grid.

HpApprDCT adheres to the same theoretical framework as [9], with key enhancements.

All compression phases described in Subsect. 3.2 have been fully parallelized by developing dedicated CUDA C kernels to run each operation efficiently.

The compression process consists of the following steps:

1. **Sub-sampling:** The kernel `sub_matrix_scalar()` implemented performs the parallel subtraction of a scalar value of 128 from the image matrix.
2. **DCT Computation:** The kernel `cuda_matrix_dct()` executes the Discrete Cosine Transform (DCT) in parallel by applying the matrix multiplication $T \cdot X_b \cdot T^T$ to each image block X_b within the image matrix X. This parallelized implementation distributes the workload across multiple threads.
3. **Quantization:** The quantization operation is executed in parallel by the `divide_matrices()` kernel, which applies element-wise division of the obtained coefficients by the predefined quantization matrix Q.

The decompression phase follows the inverse sequence of operations, with each step executed in parallel:

1. **De-quantization:** The kernel `divide_matrices()` performs the element-wise multiplication of the quantized coefficients by the predefined quantization matrix Q, restoring them to their original scale.

2. **Inverse DCT Computation:** The kernel `cuda_matrix_idct()` executes the inverse Discrete Cosine Transform (IDCT) in parallel, applying the matrix multiplication $T^T \cdot C_b \cdot T$ to each block C_b of the coefficient matrix C, reconstructing the image data.
3. **Up-sampling:** The kernel `add_matrix_scalar()` applies a parallel addition of a scalar value (128) to the reconstructed image matrix, restoring the pixel values to their original range.

To maximize computational efficiency, several optimizations have been introduced: data transfer from global to shared memory is performed in a coalesced manner to enhance memory access efficiency, the quantization matrix has been relocated to constant memory, which is optimized for read-only access, and further ones explained below.

- First, the **mapping** between image blocks and thread blocks has been **refined**. Previously, a 8×8 thread block was mapped to eight 8×8 image sub-blocks, meaning that a single thread computed an entire row of coefficients. In the new approach, an 8×8 thread block is mapped to a single 8×8 image sub-block, so that an entire row of threads computes one row of coefficients (see Fig. 3).
- The number of blocks and the number of threads per block have been carefully adjusted to **maximize** Streaming Multiprocessor (SM) **occupancy**, ensuring optimal utilization of GPU's computational resources.
- All read-only matrices have been relocated to **constant memory**, which is specifically optimized for such operations. This optimization reduces memory access latency and improves overall computational efficiency by allowing faster data retrieval compared to global memory.
- Finally, the **coalescing technique** has been employed when transferring data from global memory to shared memory (or register). This approach ensures that memory transactions are efficiently grouped, minimizing memory access overhead and maximizing memory bandwidth utilization, thereby enhancing parallel execution performance.

3.4 Implementation Details

The kernel responsible for the DCT receives the image data and the transform matrix as inputs, outputting the coefficients matrix to global memory. Within the kernel:

- Variables required for the DCT are stored in the shared memory.
- Each thread computes its global ID.
- The data are copied from global to shared memory in a coalesced manner, with each thread copying a single element.

With a one-to-one mapping between thread blocks and image blocks, each thread computes a single coefficient as shown in figure (see Fig. 3). It multiplies one

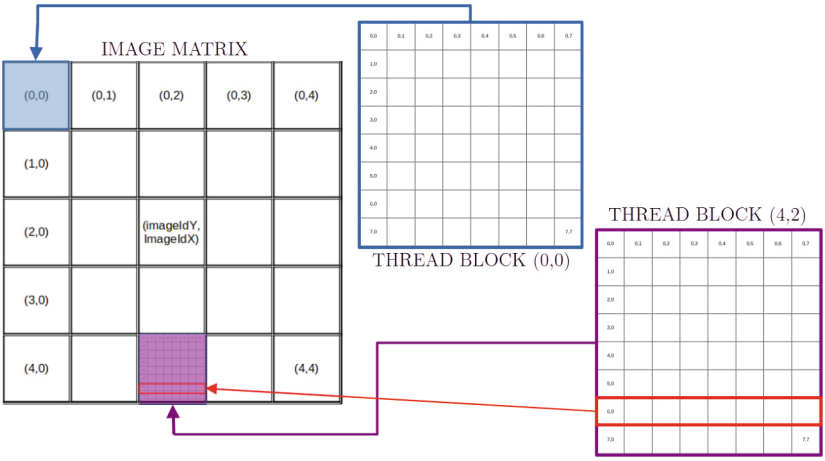


Fig. 3. Mapping between threads blocks and image blocks.

row of T by a column of the image block X_b and then multiplies the resulting value by a row of T to complete the DCT. A synchronization barrier is used between the two multiplications to ensure that intermediate results stored in `shared_matrix[]` are available to all threads.

4 Numerical Experiments

The hardware setup for the benchmarks was a Nvidia Tesla T4, featuring 2560 CUDA cores and 16 GB of GDDR6 memory. The software platform included NVIDIA driver v550.54.15 and CUDA toolkit 12.5, with the code compiled using NVCC version 12.5.82 and the parameters `-arch = sm _75, -O3`. Benchmarks were executed for all algorithms across all available input sizes.

It is worth noting that the proposed algorithm does not rely on Turing specific instructions, but exclusively utilizes CUDA C. Consequently, it is forward compatible with subsequent GPU generations. In order to run the algorithm on newer architectures simply update the compilation command by specifying the desired target architecture in the appropriate flag. (e.g., using `-arch=sm_80` for Ampere or `-arch=sm_90`). Secondly, the proposed algorithm is currently limited to single-GPU execution; scalability across multi-GPU architectures remains a key objective for future research to further enhance computational performance.

We tested four compression algorithms:

- cublasDCT: uses **cuBLAS** library, computes DCT on 8×8 sub-blocks;
- cublasDCTv2: also uses **cuBLAS**, but computes the DCT over the entire image;
- fastApprDCT: an implementation of the algorithm proposed by [9];
- HpApprDCT: the **proposed implementation**, optimized with High performance techniques.

Each assessed on six input sizes ranging from 2^8 to 2^{13}, resulting in a total of 24 runs. To ensure reliable benchmark results, each configuration was executed 100 times on randomly generated images. All execution times were recorded, and the average times were calculated.

4.1 Accuracy Evaluation

In this section, we present an extensive evaluation of the accuracy of our proposed GPU-based image compression algorithm. The analysis is performed by several significant key metrics.

First, we report the test accuracy results using the Percentage Error Energy Norm (PEEN) defined as:

$$\text{PEEN} = \sqrt{\frac{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}[I(m,n) - IR(m,n)]^2}{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} I^2(m,n)}} \times 100, \tag{2}$$

where I is the original image, and IR the reconstructed; the Mean Squared Error (MSE) defined as:

$$\text{MSE} = \frac{1}{m \cdot n}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}[I(m,n) - IR(m,n)]^2. \tag{3}$$

In addition, we compare the compression factor (the ratio between the original size and the compressed one) achieved by our method.

Table 1. Performance evaluation of HpApprDCT on the circuit image, by metrics (M) versus retained coefficients (C).

M\C	6	7	8	9	10	Standard
PEEN	9.92	8.34	8.25	8.13	6.07	4.66
MSE	79.99	56.67	55.32	53.80	30.04	17.67
Compr. Factor	1.49	1.49	1.40	1.35	1.35	1.29

In Table 1, we list the values of PEEN, MSE and the compression factor obtained by varying the number of retained coefficients. In addition, we compare the PEEN, MSE and compression factors achieved with benchmark values, to assess the efficiency of our method in preserving critical image details.

In Fig. 4, we plot PEEN values versus the number of retained coefficients (from 6 to 10) obtained by our algorithm HpApprDCT and by the PADCT algorithm proposed by [9] for comparison .

Finally, we provide visual comparisons by displaying original and compressed images of circuit and cameraman (side by side), which further highlights the practical impact of our approach on image quality.

The Figs. 5 and 6 show original cameraman and original images and the compressed one, respectively, using the standard quantization matrix.

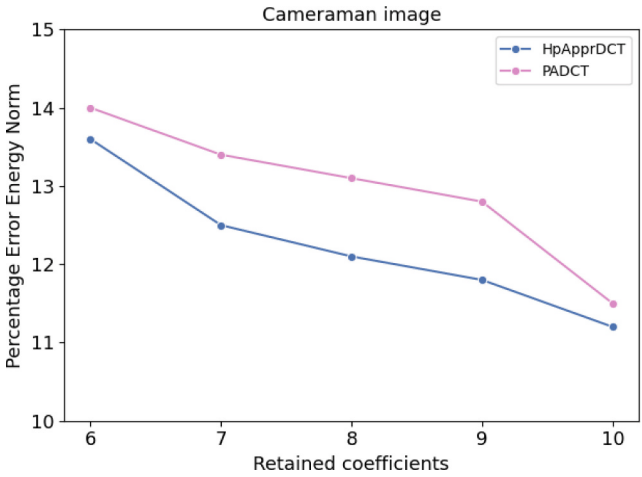


Fig. 4. PEEN values of the PADCT [9] and of the HpApprDCT versus the number of retained coeffcients, from 6 to 10.

Fig. 5. Comparison of cameraman images.

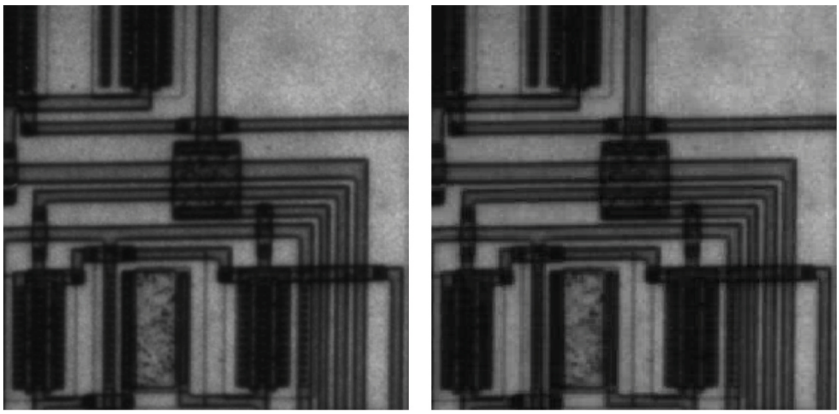


Fig. 6. Comparison of circuit images.

4.2 Efficiency Evaluation

In this section, we measure and compare the execution times of our method against alternative approaches across a range of image sizes. The evaluation is structured to include a detailed tabulation of CPU times for all algorithms, a line chart that illustrates trends in execution time, and a histogram that highlights the speedup achieved. This systematic analysis provides valuable insights into the scalability and performance benefits of our approach.

Table 2. Execution times of algorithms in milliseconds

Image size	DCT on CPU	fastApprDCT	HpApprDCT
256×256	4.7	0.28	0.07
512×512	17.9	0.33	0.12
1024×1024	72.8	0.61	0,30
2048×2048	291.7	1.65	1.04
4096×4096	1255.1	5.80	4.00
8192×8192	5005.1	20.00	14.70

To begin with, we measured the execution times of our algorithm on images of several sizes, ranging from 256 to 8192. Table 2 also shows the execution time of CPU. Next, the line chart, shown in Fig. 7, shows the trends in execution time as the image size increases, clearly illustrating how our GPU implementation scales compared to the others.

Finally, the histogram of Fig. 8 highlights the speedup achieved by the two fastest methods, providing a clear visualization of their performance gains. This

systematic analysis offers valuable insights into the scalability and overall efficiency of our approach.

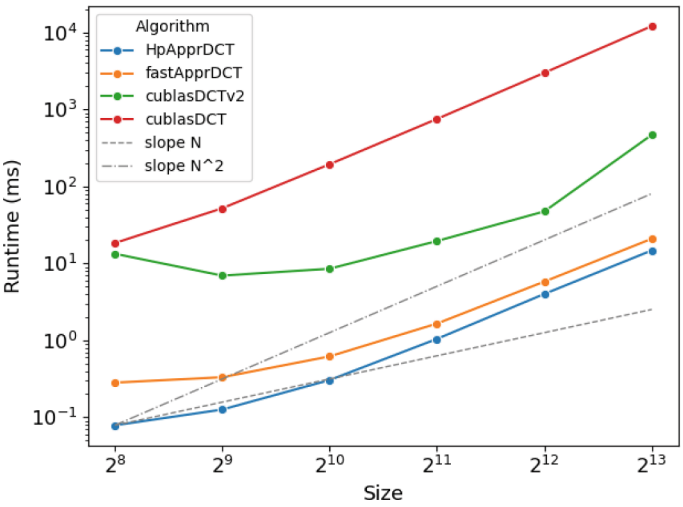


Fig. 7. Comparison of execution times.

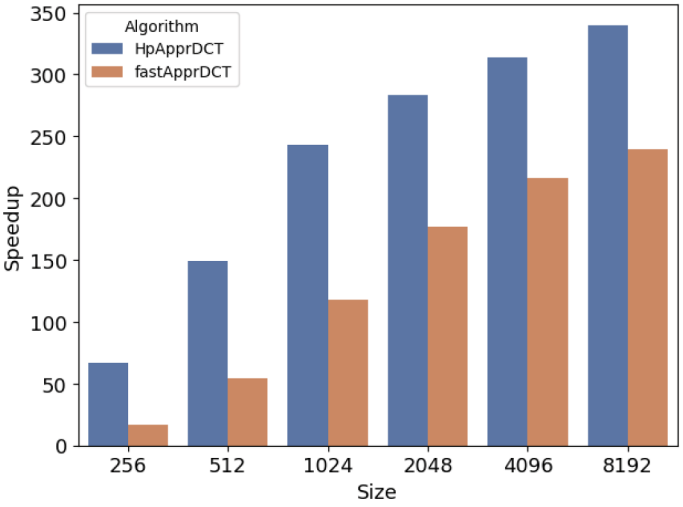


Fig. 8. Speed-up chart.

4.3 Impact of Optimization Strategies

It is important to discuss the relative impact of the individual optimization strategies based on established GPU computing literature.

Since global memory bandwidth is the primary bottleneck in parallel matrix operations [15], the use of **memory coalescing** coupled with **shared memory** is the most critical and beneficial performance enhancement. Indeed, relying on uncoalesced memory access severely bottlenecks computational throughput [18]. Secondarily, the refined **thread mapping** strategy increases the occupancy of the streaming multiprocessor (SM), which is essential to hide memory latency and fully exploit the device's resources [15]. Finally, relocating matrices to **constant memory** has a less dominant impact, but it still effectively enhances the broadcasting of read-only data across warps.

Overall, these optimizations act synergistically to achieve the excellent reported performances and ensure a highly use of hardware resources.

5 Conclusions

This study has presented an optimized implementation of the Discrete Cosine Transform (DCT) on a GPU, focusing on efficient matrix transformation techniques. The key contribution lies in the adoption of advanced algorithms that leverage memory management strategies, such as coalesced memory access and thread divergence reduction, to enhance the overall efficiency of parallel computation on a GPU.

Compared to existing solutions, the proposed algorithm (**HpApprDCT**) has demonstrated a significant performance improvement, reducing execution time by up to 31% compared to the *fastApprDCT* algorithm and more than 96% compared to *cublasDCTv2*, which relies on the cuBLAS library. Furthermore, the proposed algorithm achieves a speed-up of 340× compared to the sequential CPU version, more than 100× higher than the best speed-up obtained by *fastApprDCT*.

This performance gain enables the efficient application of DCT in scenarios where processing speed is critical, such as real-time image compression for transmission purposes.

Future developments may regard the extension the algorithm to support color image processing. As a matter of fact, the current version of the algorithm, operates exclusively on grayscale images, due to the use of the standard C library `libjpeg`., the current implementation. Further developments may regard a more extensive evaluation of the performances of the algorithm on different hardware models.

Acknowledgments. This study was partially funded by INdAM-GNCS project. A. Cardone is a member of Gruppo Nazionale per il Calcolo Scientifico, Istituto Nazionale di Alta Matematica (INdAM).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bouguezel, S., Ahmad, M.O., Swamy, M.: Low-complexity 8× 8 transform for image compression. Electron. Lett. **44**(21), 1249–1250 (2008)
2. Bouguezel, S., Ahmad, M.O., Swamy, M.: A low-complexity parametric transform for image compression. In: 2011 IEEE International Symposium of Circuits and Systems (ISCAS), pp. 2145–2148. IEEE (2011)
3. Cardone, A., De Luca, P., Galletti, A., Marcellino, L.: Solving time-fractional reaction-diffusion systems through a tensor-based parallel algorithm. Physica A Stat. Mech. Appl. **611**, 128472 (2023)
4. Ciaramella, A., De Angelis, D., De Luca, P., Marcellino, L.: Accelerated numerical simulations of a reaction-diffusion-advection model using Julia-CUDA. Mathematics **13**(9), 1488 (2025)
5. Conte, D., D'Ambrosio, R., Paternoster, B.: GPU-acceleration of waveform relaxation methods for large differential systems. Numer. Algorithms **71**(2), 293–310 (2016)
6. Cuomo, S., De Michele, P., Galletti, A., Marcellino, L.: A GPU parallel implementation of the local principal component analysis overcomplete method for DW image denoising. In: 2016 IEEE Symposium on Computers and Communication (ISCC), pp. 26–31. IEEE (2016)
7. Cuomo, S., Galletti, A., Marcellino, L.: A GPU algorithm in a distributed computing system for 3D MRI denoising. In: 2015 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), pp. 557–562. IEEE (2015)
8. Conte, D., Cuesta, E., Valentino, C.: Non-stationary wave relaxation methods for general linear systems of Volterra equations: convergence and parallel GPU implementation. Numer. Algorithms **95**(1), 149–180 (2024)
9. Haweel, R.T., El-Kilani, W.S., Ramadan, H.H.: Fast approximate DCT with GPU implementation for image compression. J. Vis. Commun. Image Represent. **40**, 357–365 (2016)
10. Haweel, T.I.: A new square wave transform based on the DCT. Sig. Process. **81**(11), 2309–2319 (2001)
11. Niemeyer, K.E., Sung, C.J.: Recent progress and challenges in exploiting graphics processors in computational fluid dynamics. J. Supercomput. **67**(2), 528–564 (2014)
12. Pennebaker, W.B., Mitchell, J.L.: JPEG: Still Image Data Compression Standard. Springer, New York (1992)
13. Pratx, G., Xing, L.: GPU computing in medical physics: a review. Med. Phys. **38**(5), 2685–2697 (2011)
14. Rea, D., Perrino, G., di Bernardo, D., Marcellino, L., Romano, D.: A GPU algorithm for tracking yeast cells in phase-contrast microscopy images. Int. J. High Perform. Comput. Appl. **33**(4), 651–659 (2019). https://doi.org/10.1177/1094342018801482
15. Ryoo, S., Rodrigues, C.I., Baghsorkhi, S.S., Stone, S.S., Kirk, D.B., Hwu, W.m.W.: Optimization principles and application performance evaluation of a multithreaded GPU using CUDA. In: Proceedings of the 13th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming, pp. 73–82. Association for Computing Machinery (2008). https://doi.org/10.1145/1345206.1345220
16. Santos, L., Magli, E., Vitulli, R., López, J.F., Sarmiento, R.: Highly-parallel GPU architecture for lossy hyperspectral image compression. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. **6**(2), 670–681 (2013)

17. Simek, V., Asn, R.R.: GPU acceleration of 2D-DWT image compression in MATLAB with CUDA. In: 2008 Second UKSIM European Symposium on Computer Modeling and Simulation, pp. 274–277. IEEE (2008)
18. Volkov, V., Demmel, J.W.: Benchmarking GPUs to tune dense linear algebra. In: Proceedings of the 2008 ACM/IEEE Conference on Supercomputing, SC '08, pp. 1–11 (2008). https://doi.org/10.1109/SC.2008.5214359

Numerical Modeling for Information Diffusion on Social Media: a Review of Recent Works

Angelamaria Cardone[1], Dajana Conte[1], Patricia Díaz de Alba[2], Samira Iscaro[1(✉)], Giovanni Pagano[3], and Beatrice Paternoster[1]

[1] Department of Mathematics, University of Salerno, Via Giovanni Paolo II 132, 84084 Fisciano, (SA), Italy
{ancardone,dajconte,siscaro,beapat}@unisa.it

[2] Department of Mathematics and Computer Science, University of Cagliari, Via Ospedale 72, 09124 Cagliari, Italy
patricia.diazda@unica.it

[3] Department of Agricultural Sciences, University of Naples Federico II, Via Università 100, 80055 Portici, (NA), Italy
giovanni.pagano5@unina.it

Abstract. Social media play a central role in information diffusion, enabling users to rapidly generate and share contents, often without adequate verification. This makes such platforms particularly susceptible to the spread of fake or partially verified information. In this work, we will present some deterministic models for the analysis of information propagation, inspired by epidemiology and formulated as systems of ordinary differential equations. Firstly, we will review the Ignorant–Spreader–Recovered (ISR) model, and some of its possible extensions that incorporate additional behavioral and structural features, including the Ignorant–Exposed–Spreader–Skeptic (IESZ) and the Ignorant–Spreader–Counter spreader–Recovered (ISCR) models. Furthermore, we will consider an age–structured version of the ISR model to better capture heterogeneity in users' activities and interactions. Firstly, we will focus on the qualitative features of these models, particularly highlighting the role of stiffness and its connection to the characteristic time scales of information spread. Secondly, since some properties of the considered models are a–priori known, as positivity, we will review the construction of adapted numerical methods, with an emphasis on the so–called Nonstandard Finite Difference (NSFD) schemes. Finally, we will show an application of the presented models and numerical approaches for the analysis of the spread of true and fake news recently shared on X.

Keywords: Information diffusion · Epidemiological models · ISR model · IESR model · IESZ model · ISCR model · age-group model · NSFD scheme · positivity preservation · stiffness analysis

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 199–220, 2027.
https://doi.org/10.1007/978-3-032-30530-5_13

1 Introduction

The rapid growth of digital technologies has reshaped the way information is produced, distributed, and consumed. In today's hyperconnected world, social media platforms have become primary channels for news dissemination, gradually surpassing traditional media such as newspapers, since they are often freely accessible and easy to use. As a result, users increasingly rely on personalized feeds, streaming content, and online communities to stay informed, contributing to a complex and dynamic information ecosystem that also includes blogs, forums, and other digital platforms.

However, alongside the benefits of accessibility and speed, this transformation has introduced significant challenges concerning the quality and reliability of information. Social media facilitate not only the spread of accurate news but also the rapid diffusion of misleading or false content. Although the dissemination of false information is not a new phenomenon, its impact has been amplified in the digital era. To better understand and address this phenomenon, recent research has focused on detection methods for false information using machine learning and AI techniques, ranging from probabilistic methods such as Bayesian classifiers, to neural network - based models that consider textual features and patterns of user interactions, as well as approaches leveraging text pre-processing and representation, including tokenization, stemming, and N-grams [18,19]. This review paper aims to present, in a structured and self-contained manner, selected recent contributions to the modeling of information diffusion and their numerical treatment, with a focus on ordinary differential equation (ODE) models for describing and predicting news dissemination, whether the information is true or false, adapted numerical methods for their solution, and strategies for data acquisition and parameter estimation, highlighting their scope of application and main limitations.

Mathematical Models

A key role in this field is played by epidemiological models, i.e. the mathematical models used to describe epidemics [27]. These models are generally of compartmental type, that is they divide the population affected by the virus in classes, as susceptibles, infected, recovered and so on, each of which ruled by a differential equation, treating the news as a virus that spreads in a population of living beings. This formulation can be done through ODEs, as e.g. in [5,7,15–17,26,32], Partial Differential Equations (PDEs) [37], as well as a stochastic formulation that incorporate probabilistic aspects of user behavior (see [32] and references therein).

Among the classical models of this type, the so–called Ignorant–Spreader–Recovered ODE model is very well known. It takes into account three groups of individuals, as the ones who can see a news and share it, called Ignorants (I), the ones who share a news called Spreaders (S), and the group of individuals who do not believe anymore in the news, called Recovered (R). There is no unique formulation of the ISR model, as several models of these kind can be found in literature. As example, in [32,33] the authors propose several ISR models in order to better catch the dynamics of information diffusion depending on whether

a news item was spread on a particular website or social media, proposing at the same time also a possible ISR model with a decay, in order to consider that a possible loss of interest towards the news may occur as time passes. Many models of interest that extend the ISR class are also mentioned in [16,32,33] and reference therein, where the authors propose an Ignorant–Exposed–Spreader–Recovered (IESR) model that takes into account the possibility that an ignorant can see a news and believe it, but without sharing anything related to it. In this case, this individual is considered as a member of the exposed class (E). This class is the counter part of the exposed class of the epidemiological models, i.e. the class of all the individuals that have been in contact with an infectious but have not develop the symptoms of the disease yet. Worth of mention are also the classes of methods of Ignorant - Spreader - Counter Spreader - Recovered (ISCR) [12,32,33] and of Ignorant - Exposed - Spreader - Skeptic (IESZ) [12,13,26]. Both of them are built in such a way that it is possible to capture more complex dynamics of the information diffusion process. In fact, the former takes into account two classes of spreaders, that is the spreaders that share a piece of news and the counter spreaders that share news with an opposite content. The latter, on the contrary, is suitable for the description of the initial phases of the information spread. It takes into account, apart from ignorants and spreaders, the classes of exposed (E), defined before, and skeptics (Z), which is the class of individuals that have seen a news but remained indifferent to it. All these models can be further refined by considering their age-structured counterparts, as done for a specific ISR model in [4,5]. In this framework, the system is formulated by accounting for the presence of individuals belonging to different age groups. Consequently, the spreading dynamics of a given piece of information may vary across age groups, reflecting different levels of interest. Each compartment of individuals is further subdivided into a set of subclasses, each corresponding to a specific age group. These subclasses retain the same epidemiological meaning and characteristics as the original compartment, while being distinguished by their associated age group index.

Furthermore, in order to more accurately characterize the dynamics of information diffusion, it is essential to investigate the fundamental properties of this typology of mathematical models. In particular, it has been rigorously established in [15,16] that the propagation speed of information is strongly associated with specific features of the underlying problem, such as stiffness. More precisely, these studies show that an increase in the stiffness of the problem described by an ISR model corresponds to a higher likelihood of information diffusion.

Numerical Methods

Regardless the model being considered, that depends on the problem we are describing, note that in all cases the solution of the ODEs system is positive. Moreover, equilibrium points could be present. Thus, these aspects should be taken into account while solving the chosen model. For this reason, positivity preserving, elementary stable and conservative numerical methods are required. Several schemes of this kind are available in literature, as e.g. modified Patankar methods [3] or numerical methods based on the use of Nonstandard Finite Dif-

ferences (NSFD) [8,9,12,28–30]. The latter class is gaining more and more attention in recent years, since NSFD methods are easy to formulate, stable and not expensive in terms of computational cost. However, NSFD schemes have order of consistency equal to one. For this reason, in recent works, as [14], also higher order NSFD schemes for general ODE problems have been formulated. In the following, we will examine these typologies of schemes for the solution of the presented models. As we will see, NSFD methods, due to their good properties, can be advantageous for a phase of parameter estimation.

Data Acquisition and Parameter Estimation

Unlike the epidemiological scenario, information diffusion models do not benefit from a data-rich framework for parameter identification. In epidemiology, parameterization is typically supported by well-structured datasets, and model parameters have a clear empirical interpretation, being directly linked to measurable biological processes. For example, clinical evidence can be used to quantify characteristic time scales of the disease, which can then be systematically translated into model rates. By contrast, such a direct correspondence between data and parameters is largely missing in information diffusion. Available datasets are often limited and unstructured and, more importantly, the parameters themselves lack a consistent interpretation across contexts, since diffusion dynamics depend strongly on the specific content and the environment in which it spreads. Several approaches have been proposed in the literature to address this goal, ranging from classical optimization-based procedures to data-driven techniques based on Machine Learning approaches. Among the latter category, a possible strategy consists in the use of Physics Informed Neural Networks (PINNs), as for example done in [31], where PINNs where used to estimate the parameters of an epidemiological model employed to describe the COVID-19 epidemic. In this work, we focus on an optimization-based approach, following [7,12,13,26,34], as it proves to be effective even when only limited data are available, such as those corresponding to the early stages of the diffusion process. In this context, after an initial data acquisition phase, a subset of the dataset is employed for parameter calibration. Starting from an initial guess of the parameters vector, the estimation procedure is formulated as a nonlinear least-squares problem. At each iteration, the ODE system is solved and the error between the model output and the observed data, restricted to the initial phase of the diffusion, is evaluated in terms of the mean squared error. This quantity is then minimized using the *lsqnonlin* MATLAB built-in routine. As we will see, although the optimization procedure requires repeated numerical solutions of the ODE system, the overall computational cost remains limited thanks to the efficiency of the NSFD schemes. This makes the approach particularly suitable for practical applications, where reliable predictions of the overall evolution of the information spread can be obtained even from a relatively small amount of early-time data.

The paper is organized as follows. Section 2 is devoted to the presentation of epidemiological models for information diffusion of ISR type, considering both the classical and age-structured formulation, while Sect. 3 is focused on more

advanced models recently proposed in literature. In Sect. 4 an in depth analysis of some key features of the above mentioned models is carried out, with a particular focus on the relation between stiffness and the velocity of news propagation. Section 5 describes the data acquisition process and parameter estimation, both fundamentals for model validation. Section 6 is dedicated to the presentation and description of NSFD schemes, while Sect. 7 present some numerical experiments related to the diffusion of true or fake news shared on X (formerly Twitter) in recent years.

2 Epidemiological Models for Information Diffusion

Epidemiological-like models have recently gained a central role in the process of information diffusion, since they allow to catch several dynamics of the considered phenomenon. The idea is to treat the information shared online as a virus that spread in a population of living beings. In particular, epidemiological models are generally formulated as compartmental models, that is they divide the population under consideration into several classes, each of which including all the individuals with the same features. Thus, according to the epidemics being described, several classes can be observed. Among the foundational models in this class, the Susceptible–Infectious–Recovered (SIR) framework proposed by Kermack and McKendrick in 1927 remains one of the most widely adopted [27,34]. It provides a division of the population in the three classes: Susceptibles, i.e. the healthy individuals that can contract a virus; Infectious, i.e. the individuals who are ill and can spread the virus; Recovered, i.e. the individuals who have recovered from the disease and have gained immunity towards it. Similarly, a model for information diffusion can be formulated. That is, the individuals who surf the net or use a social network can be divided in classes according to their status during the information spreading process. In particular, this section will be devoted to the counter part of the aforementioned SIR model for epidemics. Specifically, the Ignorant–Spreader–Recovered (ISR) model and its variants will be introduced.

2.1 ISR Models

The ISR model provides a division of the population in the following classes:

- the ignorant class $I(t)$, that reunites those who are unaware of the news under consideration, and that is the counter part of the susceptible class of the SIR model;
- the spreader class $S(t)$, that is composed of all the individuals that, believing the news in question, decide to reshare it. It is the counter part of the infectious class of the SIR model;
- the recovered class $R(t)$, composed of those who have lost interest in the previously shared news and that, as a consequence, stop sharing it. It is the counter part of the recovered class of the SIR model.

Several approaches can be adopted to formulate this model. Of particular interest is the ISR model introduced in [15]:

$$\begin{cases} I'(t) = -\bar{\beta} I(t) S(t), \\ S'(t) = \bar{\beta} I(t) S(t) - \bar{\alpha} S(t), \\ R'(t) = \bar{\alpha} S(t). \end{cases} \tag{2.1}$$

Let us consider $N(t) = S(t) + I(t) + R(t)$, i.e. the total population size at time t. Note that $N'(t) = 0$, that is, the total population size is constant, since no entries or exits from the population are taken into account. The involved parameters $\bar{\beta}$ and $\bar{\alpha}$, respectively, define the contact rate between ignorants and spreaders that rules the transitions from the ignorant class to the spreader class and the rate of transition from the spreader class to the recovered class. Specifically, these parameters are defined as:

$$\bar{\beta} = \frac{\text{INT}}{10} \qquad \text{and} \qquad \bar{\alpha} = \frac{\text{HDI}}{100},$$

where INT stands for *Internet Penetration Index* and is a parameter representing the ease of access to the Internet in a country, while HDI stands for *Human Development Index* and indicates the social development rate of the country where the news is spreading. These indexes, in particular, are recorded per country, as reported in [23,24].

The dynamics among different classes can be also described through a graph. For the ISR model, the graph is in the following Fig. 1.

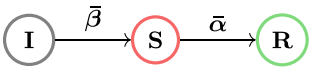


Fig. 1. Graph of the ISR model (2.1).

The aforementioned ISR model has been addressed by the authors for the analysis of fake news on social media. However, the ISR model can be formulated also for the process of information dissemination, regardless the truth of the news being shared. Interesting formulation is the one presented in [32,33] and analyzed in depth in [7]. It is an ISR model specific for social media, formulated taking into account the following parameters:

- κ, the average number of connections among users of a social network;
- β, the per capita news spreading rate per unit of time, that rules transition from the ignorant class to the spreader class;
- γ, the recovery rate, that rules the transition from the spreader class to the recovered class.

Note that the role of the parameter κ is fundamental. In fact, contrary to the epidemics, on social media encounters among individuals happen more frequently

since an encounter happens whenever people get in touch trough a reaction or a comment to a given post. This makes the process of transition between the spreader class to the recovered class much more faster than the corresponding transition from susceptibles to infectious in the epidemic scenario. The ISR model for social media reads as follows:

$$\begin{cases} I'(t) = -\beta\kappa I(t)S(t), \\ S'(t) = \beta\kappa I(t)S(t) - \gamma\kappa S(t)S(t), \\ R'(t) = \gamma\kappa S(t)S(t). \end{cases} \tag{2.2}$$

Observe that, unlike the ISR model (2.1), in this case transitions to the recovered class are governed by interactions between two spreaders. When one spreader encounters the news shared by another, it may realize that the information is false or, more generally, may lose interest in it. In both cases, this leads to a transition from the spreader class to the recovered class. This transition is mathematically expressed by the term $\gamma\kappa S(t)S(t)$.

2.2 Age-Structured ISR Models

As aforementioned, a possible approach to improve a model for information diffusion consists in considering an age-structured formulation. Such model has been proposed in [4,5] and summarized below.

Let assume that the population is divided into the n age groups. For each $i-$th group, we distinguish corresponding number of spreaders S_i, ignorants I_i and recovered R_i. The ISR model which describes the diffusion of information, proposed in [4,5], is:

$$\begin{cases} I_i'(t) = -\beta \displaystyle\sum_{j=1}^{n} M_{ij} \frac{S_j(t)}{N_j} I_i(t), \\ S_i'(t) = \beta \displaystyle\sum_{j=1}^{n} M_{ij} \frac{S_j(t)}{N_j} I_i(t) - \gamma S_i(t), \qquad t \in [0,T], \\ R_i'(t) = \gamma S_i(t), \end{cases} \tag{2.3}$$

with $i = 1, \ldots, n$. The contact rate $\beta > 0$ and the recovery rate $\gamma > 0$ are assumed to be uniform in the population groups; N_i is the (constant) number of individuals belonging to the age group g_i, $i = 1, \ldots, n$. The contacts among the spreaders and the ignorant of the different groups are modelled by the contact matrix M, which is nonnegative. In particular, in the diffusion of information in social media, M_{ij}/N_j stands for the probability of contact between an individual of the age-group i and another one of the age-group j. A practical and easy way to compute these values consist of setting $M_{ij}/N_i = p_j$, where p_j is the fraction of individuals of the $j-$th group.

In [5] it has been proved that the solution is nonnegative, for nonegative initial data, the total population is preserved, the reproduction number was

computed and is equal to $R_0 = \beta\rho(M)/\gamma$, the final size of the epidemic was derived. Moreover, a suitable NSFD method was formulated, and a Modified Patankar–Runge–Kutta scheme was adopted, both ensuring the preservation of positivity in the numerical solution.

3 Advanced Models for Information Diffusion

In the previous section, the ISR model class has been introduced, with a focus also on a possible age-structured formulation. However, in some cases, the ISR model cannot catch all the details of the scenario being described. For example, it may happen, in particular in the initial phases of a news spread, that not all the ignorant individuals that read a news are convinced of its content. Some people can be interested in the news, but need their time before sharing it. Others, on the contrary, can remain totally indifferent. Moreover, in the particular case of fake news dissemination, not all the spreaders share fake content. It can happen that a group of individuals, knowing that the shared news is fake, tries to restore the truth by sharing a news with an opposite content. Thus, to better capture these specific situations, more elaborate models extending the ISR framework can be introduced: the IESR model described in [16], the IESZ model described in [13,26] and the ISCR model proposed in [12].

3.1 IESR Model

The Ignorant - Exposed - Spreader - Recovered (IESR) model is one of the possible direct extension of the ISR model. Following the corresponding model for epidemics (see for e.g. [26]), it introduces the class of Exposed individuals E, i.e., those who have seen a piece of news and believed it but have not yet shared anything related to it yet. In particular, following [16], we present a more general IESR model, with a non constant population, in order to take into account that new individuals can enter the population, for example because influenced by oral confrontation or by news on the same topic spread through other means of communication. The new add rate of individuals to the ignorant population is ruled by the parameter Λ. Note that according to this model, when ignorant individuals get in touch with a news shared by a spreader, they believe in the news transiting into the expoded class, at rate $\tilde{\beta}$, and remaining there until the moment when they share something related to the news or when they definitely lose interest towards it. In the first case a transition into the spreader class, ruled by the parameter $\tilde{\gamma}$ can happen, while in the second case a transition into recovered class, at rate $\tilde{\delta}$, is possible. Note that also a transition from spreader to recovered is possible, ruled by the parameter $\tilde{\alpha}$ and determined by the time during which a piece of news remains trending among the most widespread news

(Fig. 2). Thus, the IESR model reads as follows:

$$\begin{cases} I'(t) = \Lambda - \tilde{\beta}\, I(t)S(t), \\ E'(t) = \tilde{\beta}\, I(t)S(t) - (\tilde{\gamma} + \tilde{\delta})E(t), \\ S'(t) = \tilde{\gamma}\, E(t) - \tilde{\alpha} S(t), \\ R'(t) = \tilde{\alpha}\, S(t) + \tilde{\delta} E(t). \end{cases} \tag{3.4}$$

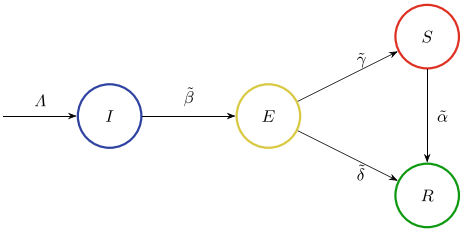


Fig. 2. Graph of the IESR model (3.4).

3.2 IESZ Model

The Ignorant–Exposed–Spreader–Skeptic model (IESZ), introduced in [13,26], slightly modifies the ISR model by considering, apart from ignorants and spreaders, the classes of Exposed E and Skeptics Z. It provides a distinction between the individuals who have seen a news, believed it but that have not shared anything related to it yet that transit into the exposed class, and the individuals who remain indifferent to the news. The latter ones are the individuals who will transit into the skeptic class. Note that, contrary to the ISR or IESR model, this one does not provide an exit from the spreader class, as clearly visible from the graph in Fig. 3. Once becoming spreader, an individual will remain in this status forever. Indeed, this model is particularly suitable to describe the initial phases of the epidemic, when we are interested in knowing how fast the number of spreaders will increase, without being interested in possible losses of interest towards the news.

The IESZ model reads as follows:

$$\begin{cases} I'(t) = -\hat{\beta} I(t) \frac{S(t)}{N} - bI(t) \frac{Z(t)}{N}, \\ E'(t) = (1-p)\hat{\beta} I(t) \frac{S(t)}{N} + (1-l)bI(t) \frac{Z(t)}{N} - \rho E(t) \frac{S(t)}{N} - \epsilon E(t), \\ S'(t) = p\beta I(t) \frac{S(t)}{N} + \rho E(t) \frac{S(t)}{N} + \epsilon E(t), \\ Z'(t) = lbI(t) \frac{Z(t)}{N}. \end{cases} \tag{3.5}$$

The model parameters are defined as follows: $\hat{\beta}$ denotes the rate at which ignorant individuals come into contact with the news and transition either to the

spreader class with probability p or to the exposed class with probability $1-p$; b denotes the rate at which ignorant individuals transition either to the skeptic class with probability l or to the exposed class with probability $1-l$. The parameter ρ represents the contact rate governing transitions from the exposed to the spreader class due to interactions with spreaders, while ϵ denotes the transition rate from exposed to spreader driven by an internal decision process, independent of further contacts. In particular, the relations among the different classes are described by the graph in Fig. 3.

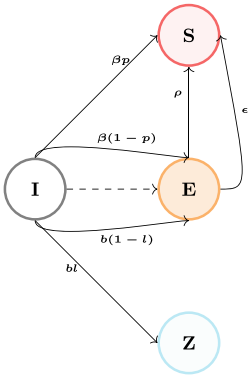


Fig. 3. Graph of the ISEZ model in (3.5).

3.3 ISCR Model

The Ignorant - Spreader - Counter spreader - Recovered (ISCR) model, firstly introduced in [32,33], is another direct extension of the ISR type models, and in particular of the ISR for social media described in the previous Sect. 2. This model is suitable to describe situations in which there are two groups of spreaders: the Spreaders S that share a news and the Counter spreaders C that try to stop the news spread by sharing content with opposite topic. Particularly interesting is the new ISCR model proposed in [12]. By considering, as done for the ISR model (2.2), κ the average number of connections of individuals, β the transition rate from the ignorant class to the spreader class and γ the transition rate from the spreader class to the recovered class, there the authors propose an ISCR model of the following form, whose corresponding graph is given by Fig. 4:

$$\begin{cases} I'(t) = \Lambda - \beta\kappa I(t)S(t) - \alpha\kappa I(t)C(t) - \delta I(t), \\ S'(t) = \beta\kappa I(t)S(t) - (\omega_1 - \omega_2)\kappa S(t)C(t) - \gamma\kappa S(t)[S(t) + C(t)] - \delta S(t), \\ C'(t) = \alpha\kappa I(t)C(t) + (\omega_1 - \omega_2)\kappa S(t)C(t) - \mu\kappa C(t)[S(t) + C(t)] - \delta C(t), \\ R'(t) = \gamma\kappa S(t)[S(t) + C(t)] + \mu\kappa C(t)[S(t) + C(t)] - \delta R(t). \end{cases} \tag{3.6}$$

The parameters α and μ are the counterparts of β and γ referred to counter spreaders, while ω_1 and ω_2 take into account the relation between spreaders and counter spreaders and rule the transitions of individuals between these classes. Furthermore, two additional parameters are considered: Λ, that as for the IESR model (3.4), is the the new addition rate to the ignorant class, and δ that has a similar meaning of the death rate for epidemics. In this context, this parameter has been introduced to take into account that during the process of information spreading some individuals might have problems accessing their accounts because reported as fake profiles or due to technical issues. Thus, the resulting model provides a total population size that evolve according to the simplified logistic equation $N'(t) = \Lambda - \delta N(t)$, where $N(t)$ indicates the total population size.

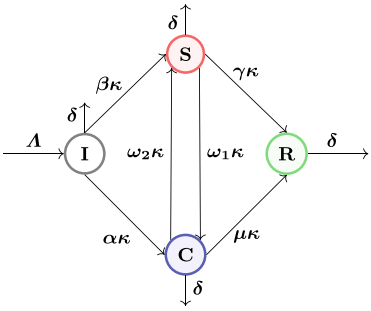


Fig. 4. Graph of the ISCR model (3.6).

4 Stiffness Analysis for the Prediction of Information Diffusion

By analyzing the structural properties of the ODE system underlying the given mathematical model, further insights into the mechanisms governing news diffusion can be obtained. In particular, as stressed in [15–17], there exist a strict link between the stiffness ratio of the system and the velocity of news dissemination, particularly in the case of fake news spread. The stiffness ratio is defined as the ratio in modulus between the largest and smallest nonzero eigenvalues of the Jacobian matrix associated with the linearized system corresponding to the ODE_S system of the model [25].

A large stiffness ratio indicates that fast and slow dynamics coexist, making the system *stiff*. In the particular case of the ISR model as explained in [15], a higher stiffness ratio corresponds to a faster transition of fake news across the population. This behavior is explained by the presence of a dominant fast time scale, which drives a rapid increase and subsequent decline in the number of spreaders. The same mechanism also leads to a faster re-establishment of the

truth, suggesting that stiffness captures not only the intensity of diffusion but also the system's responsiveness.

For example, we show below how to compute the stiffness ratio related to the ISR model (2.1). Suppose to have the following vector of initial conditions: $[I_0, S_0, R_0]^T = [I(0), S(0), R(0)]^T$. By computing the linearized system, the following Jacobian matrix is obtained:

$$J_{\alpha,\beta}(I_0, S_0) = \begin{bmatrix} -\bar{\beta}S_0 & \bar{\beta}I_0 & 0 \\ \bar{\beta}S_0 & \bar{\beta}I_0 - \bar{\alpha} & 0 \\ 0 & \bar{\alpha} & 0 \end{bmatrix}.$$

By computing the eigenvalues of this matrix, it is possible to obtain the stiffness ratio: the larger this value, the more the news will spread, affecting a high number of users. In this sense, the stiffness ratio provides a quantitative link between the mathematical structure of compartmental models and the qualitative behavior of information diffusion. It captures not only the speed of propagation, but also the efficiency of the mechanisms leading to the re-establishment of the truth, thus offering a valuable tool for the analysis of misinformation dynamics.

5 Data Acquisition and Parameter Estimation

The previous sections have introduced several mathematical models for the description of information diffusion, and also a link between mathematical properties of the model and the evolution of the phenomenon. However, in order to practically analyze this type of phenomena, it is fundamental to first construct a proper dataset to work with. In fact, there exist some tools to extract data from social networks, as for example the Python data scraper *Tweepy* or the X (formerly Twitter) portal *Twitter Developer Option*. However, the extracted data come in a qualitative format and are referred to a chosed "hashtag". In particular, the extracted information is related to user ID of the user who has shared the tweet, the date and time of the tweet and the written message. Thus, after the phase of data collecting, it is fundamental to organize the information in a quantitative way in order to construct a numerical dataset. Data are organized according to the following Algorithm 1, proposed in [7]. The algorithm assigns each user to the ignorant class until time t_0, when the first tweet is posted. At that point, the user moves into the spreader class and remains there until time $t_1 + t_{\text{exit}}$, that is, a duration t_{exit} after the moment of the last tweet t_1. The time t_{exit} needed to transition between compartments depends on the specific news being tracked. Therefore, when constructing the dataset, it is necessary to account for different transition times from the spreader to the recovered state.

Once constructed the dataset, other fundamental critical phase to be carried out for the model validation process is represented by the computation of the model parameters. Indeed, even if for particular models as the ISR given by (2.1) and presented in [15], parameters are computed starting from given tabulated indexes, generally the computation of these parameters is not trivial. On the

Algorithm 1. Constructing I, S, R vectors from Twitter data

```
I ← Ignorant Population
S ← Spreader Population
R ← Recovered Population
t ← First Data of dataset
t_end ← Last Data of dataset
t_exit ← Time to pass from S to R
for each user do
    t_0 ← first time this user tweeted
    t_1 ← last time this user tweeted
end for
loop
if t ≤ t_0 then
    I ← I + 1
    t ← t + t_exit
    go to loop
else
    if t_0 ≤ t ≤ t_1 + t_exit then
        S ← S + 1
        t ← t + t_exit
        go to loop
    else
        if t_1 + t_exit ≤ t ≤ t_end then
            R ← R + 1
            t ← t + t_exit
            go to loop
        else
            close loop
        end if
    end if
end if
```

contrary of the epidemiological scenarios in which for example parameters can be estimated by knowing some properties of the virus as its incubation period or the response of infected to medical treatments, with news dissemination this is not possible. Thus, in order to estimate the model parameters, a parameter estimation strategy is necessary. Following what has been proposed in [7,12,13], it is possible to determine a set of optimized parameters for the chosen model by minimizing a given error function. A possible choice for the latter consists in considering the mean squared error between real data at each time step t_j and the corresponding real values. This error, in particular, can be computed by considering the error on the actual number of spreaders and the number of the recovered at a given time step, in order to take into account the individuals who are spreading the considered news and the individuals who will not be involved in the news dissemination anymore. In particular, the error function reads as

follows:

$$f(param) = \sum_{j=1}^{n} (S_j - S(t_j))^2 + \sum_{j=1}^{n} (R_j - R(t_j))^2. \tag{5.7}$$

Here, n is the total number of data taken in consideration and *param* is the set of parameters to be optimized. From a computational point of view, several tools are available for the process of parameter optimization. In particular, in order to minimize the mean square error described by (5.7), it is possible to make use of the MATLAB built in function *lsqnonlin*. This routine allows the user to select the optimization algorithm used in the estimation process. Among the available options there is the trust-region reflective method, which supports parameter bounds and thus prevents solutions from taking on non-physical values, such as negative ones.

6 NSFD Schemes

In the previous section, we introduced a strategy for data acquisition and parameter estimation. As will be clear in the numerical experiments section, the approach we adopt consists of optimizing the model parameters by minimizing the error over a given set of real data. However, this procedure requires solving the system repeatedly, which makes the use of computationally expensive numerical methods impractical. Therefore, we need numerical schemes that are not only stable and accurate, but also efficient and computationally affordable. Moreover, we are dealing with models whose solutions must, e.g., remain positive, since they represent the number of individuals in each compartment of the population. For this reason, it is essential to employ numerical methods that preserve, e.g., positivity.

Among the available approaches, Nonstandard Finite Difference (NSFD) schemes, originally introduced by Mickens [28–30], provide a suitable framework, especially in the context of epidemiological-type models. These methods have been widely studied in the literature due to their ability to preserve essential qualitative properties of the continuous problem, such as positivity and stability of the equilibrium points, when appropriate construction principles are followed (see, e.g., [1,2,8–10,30]). To briefly recall these principles, consider an ODE problem of the form

$$\begin{cases} y'(t) = f(t, y), \\ y(t_0) = y_0, \end{cases} \qquad t \in [t_0, t_{end}],$$

where $f : [t_0, t_{end}] \times \mathbb{R}^m \to \mathbb{R}^m$ and $y : [t_0, t_{end}] \to \mathbb{R}^m$. Let $\{t_n = t_0 + nh,\ n = 0, \ldots, N\}$ be a uniform discretization of the time interval, with $t_N = t_{end}$. Starting from a standard finite difference (SFD) scheme, an NSFD scheme can be constructed by modifying two key components. First, the first order derivative of the unknown function is approximated through the generalized discrete form

$$y'(t_n) \approx \frac{y_{n+1} - \Psi(h)\, y_n}{\Phi(h)},$$

where $y_n \approx y(t_n)$, and $\Psi(h)$ and $\Phi(h)$ are nonnegative functions, referred to as the numerator and denominator functions. They, respectively, satisfy:

$$\Psi(h) = 1 + \mathcal{O}(h), \quad \Phi(h) = h + \mathcal{O}(h^2), \quad \text{as } h \to 0.$$

Note that the classical finite difference approximation is recovered when $\Psi(h) = 1$ and $\Phi(h) = h$. Second, nonlinear terms in $f(t, y)$ are approximated in a nonlocal manner, i.e. by approximating terms as $y^2(t_n)$ as $y_n y_{n+1}$.
By appropriately applying these modifications to a given standard explicit method, it is possible to construct schemes that remain explicit while better capturing the qualitative behavior of the continuous model, such as positivity of the solution. Moreover, since the modifications to the corresponding standard method are obtained through a perturbation only of order $\mathcal{O}(h)$, the consistency is automatically ensured (see, e.g., [2]).
In the following, we illustrate the construction of an NSFD scheme by considering the age-structured ISR model (2.3) and the ISCR model (3.6) as a representative examples.

6.1 NSFD Scheme for the Age-Structured ISR Model

In [5], these nonlocal representation of the right-hand side and denominator function have been adopted

$$\beta I_j S_i \to \beta I_j^{(k)} S_i^{(k+1)}, \gamma I_i \to \gamma I_i^{(k+1)}, \qquad \phi(h, \gamma) = \frac{1 - e^{-\gamma h}}{\gamma},$$

to derive the following NSFD method for the age-group ISR model (2.3):

$$\begin{cases} \dfrac{S_i^{(k+1)} - S_i^{(k)}}{\phi(h,\gamma)} = -\beta \displaystyle\sum_{j=1}^{n} M_{ij} \frac{I_j^{(k)}}{N_j} S_i^{(k+1)}, \\ \dfrac{I_i^{(k+1)} - I_i^{(k)}}{\phi(h,\gamma)} = \beta \displaystyle\sum_{j=1}^{n} M_{ij} \frac{I_j^{(k)}}{N_j} S_i^{(k+1)} - \gamma I_i^{(k+1)}, \\ \dfrac{R_i^{(k+1)} - R_i^{(k)}}{\phi(h,\gamma)} = \gamma I_i^{(k+1)}, \end{cases} \quad i = 1, \ldots, n, \quad k = 0, \ldots, K.$$

The proposed method preserves the total population and the positivity of the solution, for any choice of the stepsize $h > 0$. To overcome the accuracy limitations of first-order convergence of this method, Modified Patankar-Runge-Kutta methods provide a viable alternative for higher-order positivity preservation. Their formulation, however, requires a specific mathematical formalism that is beyond the scope of this work; the interested reader is referred to [5].

6.2 NSFD Scheme for the ISCR Model

By applying the previously introduced NSFD construction rules, it is possible to obtain a NSFD scheme for the ISCR model (3.6):

$$\begin{cases} \frac{I_{n+1}-I_n}{\Phi(h)} &= \Lambda - \beta\kappa I_{n+1}S_n - \alpha\kappa I_{n+1}C_n - \delta I_n, \\ \frac{S_{n+1}-S_n}{\Phi(h)} &= \beta\kappa I_{n+1}S_n - (\omega_1-\omega_2)\kappa S_{n+1}C_n - \gamma\kappa S_{n+1}(S_n+C_n) - \delta S_n, \\ \frac{C_{n+1}-C_n}{\Phi(h)} &= \alpha\kappa I_{n+1}C_n + (\omega_1-\omega_2)\kappa S_{n+1}C_n - \mu\kappa C_{n+1}(S_n+C_n) - \delta C_n, \\ \frac{R_{n+1}-R_n}{\Phi(h)} &= \gamma\kappa S_{n+1}(S_n+C_n) + \mu\kappa C_{n+1}(S_n+C_n) - \delta R_n. \end{cases}$$

For implementation, it is convenient to rewrite the NSFD scheme in explicit recursive form.

$$\begin{cases} I_{n+1} &= \dfrac{(1-\delta\Phi(h))I_n + \Lambda\Phi(h)}{1+\beta\kappa\Phi(h)S_n + \alpha\kappa\Phi(h)C_n}, \\ S_{n+1} &= \dfrac{(1-\delta\Phi(h)+\beta\kappa\Phi(h)I_{n+1})S_n}{1+(\omega_1-\omega_2)\kappa\Phi(h)C_n + \gamma\kappa\Phi(h)(S_n+C_n)}, \\ C_{n+1} &= \dfrac{(1-\delta\Phi(h)+\alpha\kappa\Phi(h)I_{n+1}+(\omega_1-\omega_2)\kappa\Phi(h)S_{n+1})C_n}{1+\mu\kappa\Phi(h)(S_n+C_n)}, \\ R_{n+1} &= (1-\delta\Phi(h))R_n + \Phi(h)\,\kappa(\gamma S_{n+1}+\mu C_{n+1})(S_n+C_n). \end{cases} \tag{6.8}$$

In particular, we indicate with N_n the discrete total population, computed as $N_n = I_n + S_n + C_n + R_n$. By construction, the obtained scheme (6.8) has first-order accuracy. Moreover, it is positivity preserving if the following denominator function is chosen:

$$\Phi(h) = \frac{1-e^{-\delta h}}{\delta}.$$

Further qualitative properties of the scheme, including the preservation of equilibrium points and their stability (elementary stability), have been proved in [12].

7 Numerical Experiments

In the previous sections, we have seen some mathematical models for the analysis of information diffusion on social media, as well as a possible strategy to estimate model parameters. Furthermore, Section 6 introduces a possible approach to develop a positivity preserving numerical scheme based on NSFD schemes. As pointed out in [15,16], in the first phases of information spread, to analyze news trend, it is possible to rely on the ISR model (2.1), since the parameters α and β are tabulated. This allows us to gain a general understanding of how a piece of news spreads within a given country. However, since our goal is to present a more accurate analysis of information spread on social media, in the following, we will analyze two specific pieces of news, rather than focusing on the general trend of news diffusion per country. By making advantage of the data extraction and parameter estimation strategies previously introduced, we analyze the evolution of the spread of the following news:

- the fake news spread with hashtag #DCBlackout [7,12,13,26], related to eventual issues with communications in the city of Washington D.C. at the beginning of June 2020, due to the Black Lives Matter movement;
- the true news spread with hashtag #MashaAmini [12], referring to the death of the Kurdish woman Mahsa Amini, who was detained by Iranian police in September 2022 after being accused of improperly wearing the hijab, in violation of Iranian law.

For further case studies, it is possible to refer to the GitHub repository [20], or to the numerical experiments sections of [7,11,12]. In this context, when a piece of news is disseminated, the focus is often not on the evolution itself, but on the identification of its peak diffusion. This corresponds to the time at which the spreading process reaches its maximum intensity, both in terms of the number of spreaders and the temporal location of this maximum. Such information is particularly crucial in the case of misinformation, as it enables the timely deployment of countermeasures to mitigate or even prevent further propagation. On the other hand, for verified or reliable news, identifying the peak can provide significant practical benefits. For example, when the information is related to a specific product, companies can leverage this insight to detect periods of highest public interest and accordingly optimize the timing of targeted marketing strategies. For these reasons, by employing the data acquisition and parameter estimation strategy as described in Section 5, particularly minimizing the error function (5.7), we compare the performances of the ISR model for social media (2.2) and of the ISCR model (3.6). In particular, we consider just a part of the dataset for the optimization strategy, that is the one with green $*$, while the remaining part of the dataset (in red $\bullet$) is used just to check how good predictions with optimmized parameters are. Note that, in the estimation procedure, all model parameters are optimized, with the exception of the add rate Λ and the mortality rate δ of the ISCR model. These two parameters are fixed consistently with the literature, specifically $\Lambda = 0.00023$ and $\delta = 0.001$, as suggested in [12,16]. Moreover, for both the considered models we assume $\kappa = \frac{\kappa_0}{N_0}$, in order to account for the dependence of the interaction rate on the population size. From a numerical perspective, the new ISCR model is repeatedly solved during the parameter estimation process by means of the NSFD scheme (6.8) introduced in Section 6. This approach presents a significant advantage, as it allows the use of relatively large time steps while still ensuring accurate results. This is particularly important since the ODE model provides only an approximation of real data: even when increasing the numerical accuracy, the error with respect to observed data does not decrease below a certain threshold. Therefore, the use of larger step sizes improves computational efficiency without compromising the overall quality of the estimation. For the ISR model, instead, we employ the MATLAB routine *ode45*. Results are displayed by Fig. 5, while optimized parameters are reported in Table 1 for the ISCR model and Table 2 for the ISR model. Moreover, Table 3 shows the errors on the total number of spreaders and on the time of the news spread peak, as well as CPU time for parameter estimation, when the ISCR model is used.

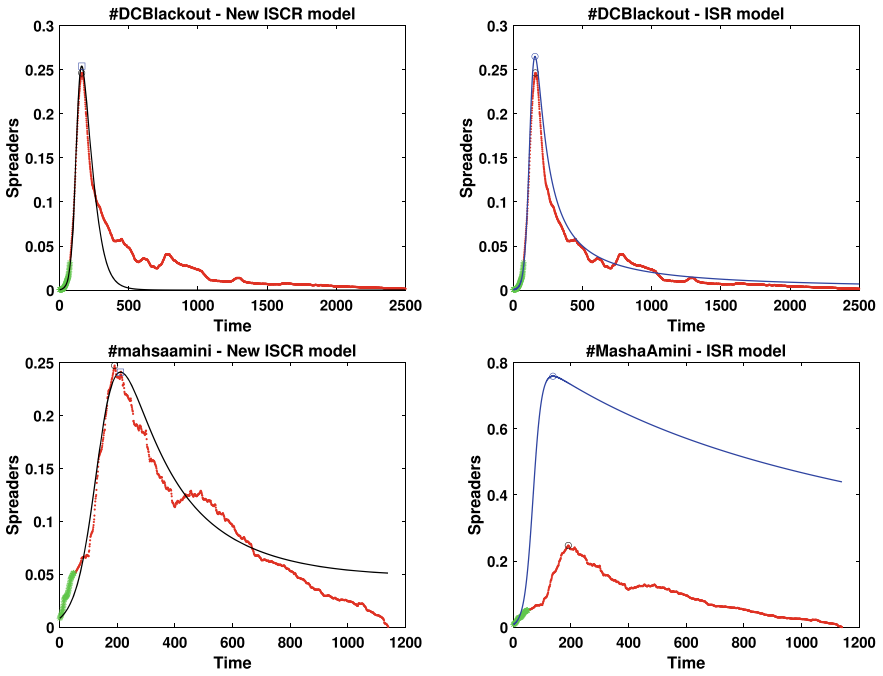


Fig. 5. News spread predictions using the ISCR model (3.6) and the ISR model (2.2), for #DCblackout (top) and #mashaamini (below). *:data for the fitting phase. •: data for prediction. -: solution of the model.

As it can be seen, the ISCR model produces better results than the ISR model on both the two case studies, since it better describes both the news evolution through time and the moment of the maximum spread of the news. This behavior is even more evident in the case of the news spread with hashtag #mashaamini, when the ISR model completely fails to catch the evolution dynamics. This outcome is not surprising, due to the main features of these two models. In fact, the ISR model provides a general description of news diffusion, whereas in the case of fake or higly debated news the ISCR model is expected to perform better, since it has been specifically designed to capture the underlying dynamics of such phenomena. Furthermore, Table 3 highlights a clear advantage of the NSFD scheme over MATLAB's ode45 solver. Indeed, even when ode45 is used with a relative tolerance of 10^{-3} to achieve errors of the same order of magnitude as the NSFD scheme (employed here with a discretization step $h = 2$), the NSFD method requires less computational time for solving the ISCR model during the parameter estimation phase. This demonstrates that NSFD achieves comparable accuracy while being more efficient. In [12], the authors also showed that the NSFD method (6.8) outperformes both the explicit and implicit Euler method, preserving the stability properties of the equilibrium points of the continuous model. Figure 5 shows also a peculiarity of the ISCR model. Although originally

Table 1. Parameter values of the ISCR model (3.6) for #DCBlackout and #mashaamini news.

News	β	γ	α	μ	k_0	ω_1	ω_2
#DCBlackout	0.30	0.03	0.30	0.04	0.20	1.66	1.05
#mashaamini	0.10	0.10	0.10	0.30	0.30	0.50	0.40

Table 2. Parameter values of the ISR model (2.2) for #DCBlackout and #mashaamini news.

News	β	γ	k_0
#DCBlackout	0.30	0.30	0.20
#mashaamini	2.02	0.03	0.03

Table 3. Comparison between NSFD and ode45 methods in terms of ErrTp, ErrSp, and CPUtime for the parameter estimation process, for the #DCBlackout and #mashaamini news.

News	**Method**	**ISCR**			**ISR**		
		ErrTp	**ErrSp**	**CPUtime**	**ErrTp**	**ErrSp**	**CPUtime**
#DCBlackout	NSFD ($h = 2$)	0.046	0.050	2.2188	0.046	0.050	2.2188
	ode45	0.045	0.050	3.5781	0.045	0.050	3.5781
#mashaamini	NSFD ($h = 2$)	0.082	0.015	0.71875	0.082	0.015	0.71875
	ode45	0.082	0.015	0.85938	0.082	0.015	0.85938

developed to study the propagation of fake news, where spreaders disseminate misinformation while counter spreaders attempt to restore the truth, the model is also suitable for the description of true news spread, as e.g. for the case of the news spread with the hashtag #mashaamini. This happens in the case of true but highly debated news, where distinct groups of individuals actively share and discuss content related to the topic.

8 Conclusion

In this work, several epidemiological models for analyzing information diffusion on social networks have been reviewed, ranging from basic frameworks such as the ISR class to more advanced formulations like the IESZ and ISCR models. In addition, a data acquisition and parameter estimation strategy has been discussed, highlighting its crucial role in the model validation phase. The importance of developing numerical methods that are both problem-oriented and computationally efficient has also been emphasized. Finally, numerical experiments have been presented, in which the proposed approach is applied to analyze the spread of news recently shared on X.

Several directions for future research are possible. First, the dataset construction process could be improved. Currently, without analyzing the content of individual tweets, it is only possible to construct the classes of ignorants, spreaders, and recovered users, while other classes, such as counter spreaders, cannot yet be defined. Furthermore, the number of data points used for parameter optimization (the green data) is currently selected empirically based on the news dynamics. Future work could aim to systematically link the number of points used for optimization phase to the total number of available data. Second, regarding the numerical method, it would be interesting to explore the application of positivity-preserving NSFD schemes of high order, following recent developments in the literature [14], to improve both accuracy and efficiency.

Acknowledgments. This study was partially funded by the Italian Ministry of University and Research (MUR), through: PRIN 2022 project (CUP: E53D23017940001) DESTINY (Stochastic Numerical Modelling for Sustainable Innovation), granted within the framework of the call relating to the scrolling of the final rankings; INdAM-GNCS projects. All authors are members of Gruppo Nazionale per il Calcolo Scientifico, Istituto Nazionale di Alta Matematica (INdAM).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Anguelov, R., Dumont, Y., Lubuma, J.S., Shillor, M.: Dynamically consistent nonstandard finite difference schemes for epidemiological models. J. Comput. Appl. Math. **255**, 161–182 (2014)
2. Anguelov, R., Lubuma, J.M.S.: Contributions to the mathematics of the nonstandard finite difference method and applications. Numer. Methods Partial Differ. Equ. **17**(5), 518–543 (2001)
3. Blanes, S., Iserles, A., Macnamara, S.: Positivity-preserving methods for ordinary differential equations. ESAIM Math. Model Numer. Anal. **56**(6), 1843–1870 (2022)
4. Cardone, A., Díaz de Alba, P., Paternoster, B.: Influence of age group in the spreading of fake news: Contact matrices in social media. In: 2022 16th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), (pp. 515–521). IEEE (2022)
5. Cardone, A., Díaz de Alba, P., Paternoster, B.: Analytical properties and numerical preservation of an age-group susceptible-infected-recovered model: application to the diffusion of information. J. Comput. Nonlinear Dyn. **19**(6), 061006 (2024)
6. Casillo, M., et al: A multi-feature bayesian approach for fake news detection. In: International Conference on Computational Data and Social Networks Cham: Springer International Publishing 2020. pp. 333–344 (2020)
7. Castiello, M., Conte, D., Iscaro, S.: Using epidemiological models to predict the spread of information on Twitter. Algorithms **16**(8), 391 (2023)
8. Conte, D., D'Ambrosio, R., Giordano, G., Mottola, S., Paternoster, B.: Stiiff problems nowadays: Novel numerics and fake news. In: AIP Conference Proceedings (Vol. 2849, No. 1, p. 020004). AIP Publishing LLC (2023)

9. Conte, D., Guarino, N., Pagano, G., Paternoster, B.: Positivity-preserving and elementary stable nonstandard method for a COVID-19 SIR model. Dolomites Res. Notes Approx. **15**(5), 65–77 (2022)
10. Conte, D., Guarino, N., Pagano, G., Paternoster, B.: On the advantages of nonstandard finite difference discretizations for differential problems. Numer. Anal. Appl. **15**(3), 219–235 (2022)
11. Conte, D., Iscaro, S., Pagano, G.: On the use of Nonstandard Finite Difference methods for the analysis of information diffusion. Submitted
12. Conte, D., Iscaro, S., Pagano, G., Paternoster, B. Adapted numerical modeling for fake news diffusion on social networks. J. Comput. Appl. Math., 117196 (2025)
13. Conte, D., Iscaro, S., Paternoster, B.: Predicting information diffusion on social media using an epidemiological approach. In: International Conference on Computational Science and Its Applications, (pp. 208-224). Cham: Springer Nature Switzerland (2025)
14. Conte, D., Pagano, G., Roldán, T.: High order nonstandard finite-difference methods. Appl. Math. Comput. **510**, 129681 (2026)
15. D'Ambrosio, R., Giordano, G., Mottola, S., Paternoster, B.: Stiffness analysis to predict the spread out of fake information. Future Internet **13**(9), 222 (2021)
16. D'Ambrosio, R., Díaz de Alba, P., Giordano, G., Paternoster, B.: A modified seir model: stiffness analysis and application to the diffusion of fake news. In: International Conference On Computational Science and its Applications. Cham: Springer International Publishing, pp. 90–103 (2022)
17. D'Ambrosio, R., Mottola, S., Paternoster, B.: Stiffness ratio and the diffusion of fake news. In: American Institute of Physics Conference Series (Vol. 2425, No. 1, p. 090004) (2022)
18. Giordano G., Mottola S., Paternoster B.: A short review of some mathematical methods to detect fake news. Int. J. Circ., Syst. Signal Process (2020)
19. Giordano G., Mottola S., Paternoster B.: Some mathematical aspects to detect fake news: a short review. In: Proceedings - 2nd International Conference on Mathematics and Computers in Science and Engineering, MACISE 2020, (2020)
20. GitHub repository: https://github.com/siscaro/Codes-for-the-Analysis-of-Information-Diffusion-via-ISR-and-ISCR-models. Accessed 2026 Apr 30
21. Hoang, M.T., Ehrhardt, M.: A second-order nonstandard finite difference method for a general Rosenzweig - MacArthur predator-prey model. J. Comput. Appl. Math. **444**, 115752 (2024)
22. Hoang, M.T., Pham, H.T.: Global dynamics and numerical simulation of a modified epidemiological model for viral marketing on social networks. Math. Comput. Simul. **228**, 225–244 (2025)
23. United Nations Development Programme, Human Development Data Center: http://hdr.undp.org/en/data. Accessed Oct 13 2024
24. Wikipedia, List of countries by number of Internet users. https://en.wikipedia.org/wiki/List_of_countries_by_number_of_Internet_users. Accessed Oct 13 2024
25. Lambert, J.D.: Numerical methods for ordinary differential systems. (Vol. 146). New York: Wiley (1991)
26. Maleki, M., Mead, E., Arani, M., Agarwal, N.: Using an epidemiological model to study the spread of misinformation during the black lives matter movement. Preprint arXiv:2103.12191 (2021)
27. Martcheva, M.: An introduction to mathematical epidemiology. New York, Springer **61**, 9–31 (2015)
28. Mickens, R.E.: Nonstandard Finite Difference Models of Differential Equations. World Scientific (1993)

29. Mickens, R.E.: Dynamic consistency: a fundamental principle for constructing nonstandard finite difference schemes for differential equations. J. Diff. Eq. Appl. **11**, 645–653 (2005)
30. Mickens, R.E.: Calculation of denominator functions for nonstandard finite difference schemes for differential equations satisfying a positivity condition. Numer. Methods Partial Differ. Equ. **23**(3), 672–691 (2007)
31. Millevoi, C., Pasetto, D., Ferronato, M.: A Physics-Informed Neural Network approach for compartmental epidemiological models. PLoS Comput. Biol. **20**(9) (2024)
32. Muhlmeyer, M., Agarwal, S.: Information spread in a social media age: modeling and control. CRC Press (2021)
33. Muhlmeyer, M., Agarwal, S., Huang, A.J.: Modeling social contagion and information diffusion in complex socio-technical systems. IEEE Syst. J. **14**(4), 5187–5198 (2020)
34. Treibert, S. M.: Mathematical Modelling and Nonstandard Schemes for the Corona Virus Pandemic. Springer Spektrum (2021)
35. Twitter Developer Option. Available online: www.developer.twitter.com. Accessed on 12 July 2022
36. Tweepy Documentation. Available online: https://docs.tweepy.org/en/latest/. Accessed on 12 July 2022
37. Wang, H., Wang, F., Xu, K.: Modeling information diffusion in online social networks with partial differential equations. Springer Nature, 7 (2020)

Downsampling Strategies for LES Data from High-Fidelity CFD Solvers

Raffaella Sapio[1,2(✉)], Carmine De Lucia[2], Andrea Mastellone[2], Alfonso Matrone[2], Gerardo Toraldo[1], and Ferdinando Zullo[1]

[1] University of Campania Luigi Vanvitelli, Viale Lincoln 5, 81100 Caserta, Italy
{raffaella.sapio,gerardo.toraldo,ferdinando.zullo}@unicampania.it
[2] Centro Italiano Ricerche Aerospaziali (CIRA), Via Maiorise, 81043 Capua, Italy
{c.delucia,a.mastellone,a.matrone}@cira.it

Abstract. Downsampling is a fundamental operation in the analysis of high-fidelity Computational Fluid Dynamics (CFD) data, particularly in the context of Large-Eddy Simulation (LES), where flow fields are already represented as filtered, cell-averaged quantities. Recently, machine-learning-integrated approaches have been increasingly employed to accelerate the convergence of CFD solutions. Training neural networks in this setting requires highly accurate data, and one effective strategy is to downsample CFD solutions computed on finer grids. However, designing downsampling procedures that remain consistent with the physical and numerical structure of finite-volume solvers remains a challenging task. In this work, we investigate grid-to-grid downsampling strategies for LES data on structured meshes. We consider restriction operators that map fine-grid fields to coarser resolutions by aggregating values over blocks of cells, including arithmetic averaging, distance-weighted averaging, and extrema-based operators (minimum and maximum).

The methods are assessed through qualitative comparisons and quantitative metrics, including conservation error, blockwise discrepancies, and interface-based errors relevant to flux computations. The results show that arithmetic averaging provides the most consistent and physically meaningful coarse representation, while distance-weighted averaging offers comparable accuracy but lacks strict conservation. Extrema-based operators introduce significant distortions.

Keywords: Computational Fluid Dynamics (CFD) · Large-Eddy Simulation (LES) · Downsampling

1 Introduction

Computational Fluid Dynamics (CFD) plays a central role in the numerical analysis of fluid flows arising in engineering and physical sciences. The governing equations, typically the Navier–Stokes equations, describe the conservation of mass, momentum, and energy in a fluid and give rise to highly nonlinear and multi-scale dynamics. In turbulent regimes, the flow exhibits a wide range

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 221–238, 2027.
https://doi.org/10.1007/978-3-032-30530-5_14

of interacting spatial and temporal scales, making direct numerical simulation (DNS) prohibitively expensive for most practical applications. As a consequence, various modeling strategies have been developed to balance accuracy and computational cost, including Reynolds-Averaged Navier–Stokes (RANS) methods and Large-Eddy Simulation (LES) [11,16]. Among these approaches, LES represents a widely adopted compromise, in which the large-scale turbulent structures are explicitly resolved while the effect of the unresolved small scales is modeled. From a mathematical viewpoint, LES can be interpreted as a filtering operation applied to the Navier–Stokes equations, leading to a coarse-grained representation of the flow field [7]. In practical implementations, the computational grid and the numerical discretization act as an implicit filter, so that the stored solution corresponds to a filtered, cell-averaged field defined on the mesh. In recent years, the increasing availability of high-fidelity CFD data and advances in machine learning (ML) have opened new perspectives for the development of hybrid modeling approaches. In particular, ML techniques have been successfully employed to enhance numerical solvers, accelerate simulations, and reconstruct fine-scale features from coarse data [6,10,12]. A key aspect of these approaches is the ability to transfer information across resolutions, for instance when generating training datasets, designing surrogate models, or coupling numerical solvers with data-driven components. In this context, downsampling plays a fundamental role. Although often treated as a simple preprocessing step, downsampling in CFD is a non-trivial operation that involves reducing the spatial resolution of the data while preserving relevant physical and numerical properties. Different strategies—ranging from geometric coarsening and averaging to more complex filtering or data-driven approaches—can lead to significantly different coarse representations of the same flow field [9,14,17]. In particular, the choice of the downsampling operator directly affects the preservation of conservation laws, the representation of turbulent structures, and the consistency with the underlying numerical scheme. Despite its importance, the problem of defining suitable downsampling strategies for CFD data remains relatively underexplored, especially in the context of finite-volume LES solvers. In such settings, the solution is naturally expressed in terms of cell-averaged quantities, and numerical fluxes depend critically on interface values reconstructed from these averages. Therefore, a meaningful downsampling procedure should not only approximate the fine-grid field but also remain consistent with the conservative structure of the discretization and with the quantities entering the flux computation.

The objective of this work is to investigate and compare different downsampling strategies for CFD data generated by high-fidelity LES solvers. We focus on structured grids and consider restriction operators that map fine-grid data to coarser resolutions by aggregating values over blocks of cells. In particular, we analyze four strategies: arithmetic averaging, distance-weighted averaging, and extrema-based operators (minimum and maximum). These methods are evaluated in terms of their ability to preserve conservation properties, represent fine-scale variability, and maintain consistency with interface quantities relevant to finite-volume schemes. The analysis is carried out on a set of LES snapshots

obtained from a representative test case, and a series of quantitative metrics are introduced to assess the performance of each strategy. The results provide insight into the trade-offs between physical consistency, numerical accuracy, and feature preservation, and allow us to identify the most suitable approaches for coarse-grid representation of LES data.

The remainder of the paper is organized as follows. In Sect. 2, we provide an overview of downsampling strategies for CFD data. Section 3 describes the characteristics of the considered LES datasets. In Sect. 4, we introduce the downsampling operators and analyze their properties. Section 5 presents the numerical tests and comparative results. Finally, conclusions and perspectives for future work are discussed in Sect. 6.

2 Overview on Downsampling Strategies for CFD Data

Although often informally associated with simple subsampling, downsampling encompasses fundamentally different operations depending on the context, including mesh coarsening, filtering or coarse graining, adaptive resolution reduction, and data-driven compression. These approaches are not equivalent: each induces a distinct transformation of the underlying flow field and therefore preserves different invariants, such as conservation laws, spectral content, or turbulence statistics. Consequently, the choice of downsampling strategy must be aligned with the intended application, whether it is solver acceleration, turbulence modeling, data storage, or machine-learning-based reconstruction. A first major class of methods is geometric downsampling via mesh coarsening, in which the computational grid is explicitly reduced. This approach is central to multigrid methods and hierarchical discretizations, where coarse representations are constructed either geometrically (e.g., structured coarsening or agglomeration of elements) or algebraically through operator-based techniques. In this setting, restriction and prolongation operators define the transfer of solution variables between grid levels, and careful design is required to maintain stability and conservation properties. While mesh coarsening directly reduces the number of degrees of freedom and thus computational cost, it may degrade the representation of localized flow features such as boundary layers, shocks, or vortical structures if the coarsening is too aggressive or not adapted to the flow physics [5,17]. A second class consists of field-based downsampling through filtering or volume averaging, which is particularly relevant in the context of turbulence modeling. In large-eddy simulation, the coarse representation of the flow is interpreted as a filtered field, where unresolved scales are removed and their effects are modeled. This perspective, formalized by Germano's filtering framework, emphasizes that a physically consistent coarse field should be obtained via a low-pass filtering operation rather than pointwise subsampling [7,16]. Common filtering strategies include box filters, Gaussian filters, and spectral cut-off filters, each defining a characteristic filter width that separates resolved and unresolved scales. In contrast to implicit filtering—where the grid and numerical discretization act as the effective filter—explicit filtering allows greater control over the

scale separation and has been shown to improve robustness and consistency in certain LES formulations [8,15]. An important distinction within field-based approaches is that between decimation and averaging. Pure decimation, such as selecting every k-th grid point or cell value, is computationally inexpensive but is generally inconsistent with finite-volume data, since it discards most of the cell-averaged information within each coarse control volume. As a consequence, it may not preserve integral quantities over the coarse cells and may lead to non-physical representations of the flow. Averaging-based methods, by contrast, act as discrete low-pass filters and, on uniform grids, correspond to volume averaging, making them consistent with the finite-volume interpretation of the solution as a filtered field. This distinction has become particularly relevant in recent data-driven CFD workflows, where the generation of low-resolution training data significantly influences the performance and physical consistency of learned models [6]. A third category includes adaptive and feature-aware downsampling strategies, in which resolution is reduced non-uniformly across the domain. These methods aim to retain high resolution in dynamically important regions—such as shear layers, vortices, or discontinuities—while coarsening regions of smooth flow. Adaptive mesh refinement, multiresolution wavelet methods, and subsampling-based compression techniques fall within this class. Such approaches are especially valuable in large-scale simulations and in situ data processing, where storage and I/O constraints necessitate selective data reduction. However, adaptive strategies introduce additional complexity in terms of error estimation, refinement criteria, and the preservation of conservation properties across interfaces [9,18]. A fourth class is spectral or modal downsampling, where the flow field is represented in a reduced basis rather than on a coarser physical grid. Examples include Fourier truncation, proper orthogonal decomposition, and wavelet-based representations. These approaches are particularly effective for capturing dominant energetic structures and for reduced-order modeling, as they concentrate information in a small number of modes. However, the resulting coarse representation may not directly correspond to a standard CFD discretization, and additional reconstruction or projection steps are typically required to recover physical-space fields [1,14]. More recently, data-driven downsampling and reconstruction strategies have gained significant attention. In this setting, low-resolution flow fields are generated—typically via filtering, averaging, or coarsening—and used as inputs to machine learning models trained to reconstruct high-resolution fields or to correct coarse-grid simulations. Approaches include supervised and unsupervised super-resolution, physics-informed learning, and hybrid solvers that embed learned closures into coarse CFD computations. A key insight from this literature is that the downsampling operator is not merely a preprocessing step but a critical component that shapes the statistical structure of the data and the learnability of fine-scale features [10,12]. Overall, downsampling in CFD should be viewed as a design choice rather than a purely technical operation. Mesh-based, filtering-based, adaptive, spectral, and learned strategies each provide different trade-offs between computational efficiency, physical fidelity, and reconstructability. In particular, it is generally not

possible to simultaneously preserve conservation laws, spectral accuracy, and small-scale intermittency under aggressive resolution reduction. Therefore, the selection of a downsampling strategy must be guided by the specific objectives of the study, whether they involve numerical efficiency, turbulence-consistent modeling, data compression, or data-driven inference.

3 Considered CFD Data

The data considered in this work are obtained from high-fidelity simulations performed with a Large-Eddy Simulation solver, specifically SPARK-LES [2–4]. This solver computes the evolution of fluid flows by numerically solving the compressible Navier-Stokes equations using a finite-volume formulation on structured grids. As a consequence, the resulting datasets consist of discrete representations of physical fields defined over a set of control volumes, where conservation laws are enforced at the discrete level through flux balances across cell interfaces. More precisely, the computational domain is partitioned into a structured Cartesian mesh, and physical quantities are stored as cell-centered variables. Each cell contains averaged values of the primary flow quantities, including velocity components, pressure, density, and, depending on the physical model, additional variables such as temperature or species concentrations. These quantities represent volume-averaged states, consistent with the finite-volume discretization, rather than pointwise values. The evolution of these variables is governed by the balance of convective and diffusive fluxes computed at cell faces using high-order interpolation schemes. Since SPARK-LES operates within the LES framework, the data inherently correspond to a filtered representation of the flow field. In particular, only the large-scale turbulent structures are explicitly resolved on the computational grid, while the effects of unresolved small scales are modeled. Therefore, the grid spacing defines an effective filter width, and the stored solution can be interpreted as a coarse-grained version of the true physical flow. This intrinsic filtering plays a crucial role when further downsampling is applied, as it implies that the data are already the result of a scale-reduction process. From a numerical standpoint, the datasets exhibit several important properties that must be preserved during downsampling. First, the finite-volume formulation guarantees discrete conservation of mass, momentum, and energy, meaning that any coarsening strategy should ideally maintain consistency with these conservation laws. Second, the solution is strongly influenced by flux evaluations at cell interfaces, which depend on interpolation procedures and local flow gradients. Inaccuracies in these interface fluxes can significantly affect the overall solution quality, especially in turbulent regimes. Consequently, downsampling methods that ignore inter-cell flux consistency may introduce non-physical artifacts. Additionally, the data are typically high-dimensional and multi-scale. Even in relatively simple test cases such as the Taylor–Green vortex, the flow exhibits evolving vortical structures and a cascade of energy across scales, making it sensitive to resolution changes. In more complex configurations, such as compressible or reacting flows, the coupling between turbulence, thermodynamics,

and possibly chemical kinetics further increases the complexity of the data. This multi-scale nature implies that naive resolution reduction (e.g., simple subsampling) can lead to loss of physically relevant information, including energy spectra and coherent structures. Finally, the data are generated through explicit time integration schemes, often based on multi-stage Runge-Kutta methods, resulting in time-resolved sequences of flow fields. This temporal dimension introduces additional constraints: downsampling strategies must not only preserve spatial consistency but also avoid introducing spurious temporal dynamics when applied across successive time steps.

4 Downsampling Strategies

In this work, we consider downsampling procedures aimed at transferring solution fields from a fine computational grid to a coarser grid while preserving, as much as possible, the physical and numerical properties of the original data. The data originate from Large-Eddy Simulation computations performed with a finite-volume solver such as SPARK-LES, where the solution is represented as cell-averaged quantities over a structured mesh.

4.1 Grid Hierarchy and Notation

Let Ω^f denote the fine grid and Ω^c the corresponding coarse grid. We assume a structured hierarchy such that each coarse control volume $C \in \Omega^c$ is defined as the union of a set of non-overlapping fine-grid cells:

$$C = \bigcup_{c_f \in \mathcal{F}(C)} c_f, \tag{1}$$

where $\mathcal{F}(C) \subset \Omega^f$ denotes the set of fine cells contained in C.

Let $u^f : \Omega^f \to \mathbb{R}$ be a scalar field defined on the fine grid (e.g., pressure, density, or a component of velocity). The goal of the downsampling procedure is to construct a coarse field $u^c : \Omega^c \to \mathbb{R}$ by applying an aggregation operator over the fine cells in $\mathcal{F}(C)$.

The extension to vector-valued fields (e.g., velocity $\mathbf{u}$) is performed component-wise.

4.2 Arithmetic Averaging

The first and most natural strategy is based on arithmetic averaging. For each coarse cell $C \in \Omega^c$, we define

$$u^c(C) = \frac{1}{|\mathcal{F}(C)|} \sum_{c_f \in \mathcal{F}(C)} u^f(c_f), \tag{2}$$

where $|\mathcal{F}(C)|$ denotes the number of fine cells in C.

On uniform grids, this operation is equivalent to a discrete volume average and is therefore consistent with the finite-volume interpretation of the solution. In particular, if u^f represents a conserved quantity, the arithmetic mean preserves integral quantities across scales up to discretization error. Moreover, this operator can be interpreted as a low-pass filter, removing small-scale fluctuations while retaining the large-scale structure of the field.

For these reasons, arithmetic averaging is the most physically consistent downsampling strategy in the context of LES data.

4.3 Distance-Weighted Averaging

In addition to uniform arithmetic averaging, we consider a weighted averaging strategy in which fine-grid values closer to the center of a coarse cell are assigned a larger contribution. For each coarse cell $C \in \Omega^c$, the downsampled value is defined as

$$u^c(C) = \sum_{c_f \in \mathcal{F}(C)} w(c_f; C)\, u^f(c_f), \tag{3}$$

where the weights satisfy

$$w(c_f; C) \geq 0, \qquad \sum_{c_f \in \mathcal{F}(C)} w(c_f; C) = 1. \tag{4}$$

In the implementation considered here, the weights are chosen to be inversely proportional to the distance between the fine-cell location and the geometric center of the coarse cell. Denoting by $\mathbf{x}_{c_f}$ the position of the fine cell c_f and by $\mathbf{x}_C$ the center of the coarse cell C, we define

$$w(c_f; C) = \frac{\dfrac{1}{\|\mathbf{x}_{c_f} - \mathbf{x}_C\| + \varepsilon}}{\displaystyle\sum_{c'_f \in \mathcal{F}(C)} \frac{1}{\|\mathbf{x}_{c'_f} - \mathbf{x}_C\| + \varepsilon}}, \tag{5}$$

where $\varepsilon > 0$ is a small regularization parameter introduced to avoid singularities when a fine-cell center coincides with the coarse-cell center.

This construction yields a convex combination of the fine-cell values, so that the weighted coarse value remains bounded between the local minimum and maximum:

$$\min_{c_f \in \mathcal{F}(C)} u^f(c_f) \leq u^c(C) \leq \max_{c_f \in \mathcal{F}(C)} u^f(c_f). \tag{6}$$

Compared with arithmetic averaging, this operator gives greater importance to the central part of the coarse cell and reduced influence to values located near its boundaries. As a consequence, it may provide a representation that is more localized and less diffusive, especially when the field exhibits strong intra-cell gradients. However, unlike uniform averaging, distance-weighted averaging is generally not conservative with respect to the integral of the field over the coarse

control volume. Indeed, conservation would require weights proportional to fine-cell volumes, whereas the present choice depends on the geometric distance from the cell center. Therefore, while the weighted strategy may better preserve local features around the coarse-cell centroid, it does not in general preserve integral quantities across scales.

4.4 Maximum-Based Downsampling

The second strategy assigns to each coarse cell the maximum value among the corresponding fine cells:

$$u^c(C) = \max_{c_f \in \mathcal{F}(C)} u^f(c_f). \tag{7}$$

This operator emphasizes the most intense local features within each coarse region. It can be particularly useful for detecting strong gradients, localized peaks, or intermittent structures, which are common in turbulent flows.

However, this approach is non-conservative and introduces a systematic bias toward larger values. As a result, it does not preserve integral quantities and may distort statistical properties such as mean values or energy distributions when applied repeatedly across multiple coarsening levels.

4.5 Minimum-Based Downsampling

Similarly, a minimum-based strategy can be defined as

$$u^c(C) = \min_{c_f \in \mathcal{F}(C)} u^f(c_f). \tag{8}$$

This operator highlights the lowest values within each coarse cell and may be useful in identifying low-intensity regions or preserving lower bounds of physical quantities. As in the case of the maximum operator, this method is non-conservative and may lead to systematic underestimation of the field.

Remark 1. We underline that the same strategy is applied for the data arising from interface values.

4.6 Properties of the Downsampling Operators

The four strategies described above exhibit different mathematical and physical properties.

Conservation. Arithmetic averaging is consistent with the conservation principles underlying finite-volume methods, as it approximates volume integration over the coarse cells on uniform grids. Distance-weighted averaging, while still based on a convex combination of fine-cell values, is generally not conservative because the weights are not proportional to cell volumes. Maximum and minimum operators are also non-conservative.

Linearity. Both arithmetic averaging and distance-weighted averaging are linear operators, i.e.,

$$\mathcal{R}(\alpha u + \beta v) = \alpha \mathcal{R}(u) + \beta \mathcal{R}(v), \tag{9}$$

where $\mathcal{R}$ denotes the restriction operator. This property is important for consistency with the linear structure of the governing equations. Maximum and minimum operators are nonlinear and therefore do not satisfy this property.

Boundedness. Averaging and distance-weighted averaging satisfy

$$\min_{c_f \in \mathcal{F}(C)} u^f(c_f) \leq u^c(C) \leq \max_{c_f \in \mathcal{F}(C)} u^f(c_f), \tag{10}$$

since both are convex combinations of the fine-cell values. The maximum and minimum operators attain the upper and lower bounds, respectively.

Spectral Behavior. Arithmetic averaging acts as a discrete low-pass filter and damps high-frequency components of the field. Distance-weighted averaging also has a smoothing effect, but with a stronger emphasis on the central region of the coarse cell, thus producing a filtered field that is less uniform than the arithmetic mean. Maximum and minimum operators do not correspond to classical linear filters and may introduce distortions in the spectral content.

4.7 Implications for LES Data

Since the input data are obtained from LES simulations, they already represent a filtered version of the underlying flow. Therefore, additional downsampling introduces a second level of filtering or transformation. In this context, arithmetic averaging is consistent with the LES framework, as it preserves the interpretation of the solution as a coarse-grained field.

On the other hand, extrema-based operators (maximum and minimum) modify the physical meaning of the variables and may alter the balance between resolved and unresolved scales. Their use should therefore be carefully evaluated depending on the application, particularly in relation to turbulence statistics and energy transfer across scales.

5 Numerical Tests

5.1 Comparison Metrics

To assess the quality of the different downsampling strategies, we introduce a set of comparison metrics designed to quantify both volumetric consistency and interface accuracy. Since the original high-fidelity data are defined as cell-centered quantities on the fine grid, while flux computations in the finite-volume solver depend on interface values, it is important to evaluate the coarse representation from both perspectives.

Let u^f denote the fine-grid field and u^c the corresponding downsampled coarse-grid field. For a given coarse cell $C \in \Omega^c$, let $\mathcal{F}(C)$ be the set of fine cells contained in C, and let $N_C = |\mathcal{F}(C)|$ be their number.

Global Conservation Error. A first metric measures the discrepancy between the total fine-grid quantity and the total coarse-grid quantity:

$$Q^f = \sum_{c_f \in \Omega^f} u^f(c_f)\,\Delta V_f, \qquad Q^c = \sum_{C \in \Omega^c} u^c(C)\,\Delta V_C, \tag{11}$$

where ΔV_f and ΔV_C denote the fine- and coarse-cell volumes, respectively. The relative conservation error is then defined as

$$\varepsilon_{\text{cons}} = \frac{|Q^c - Q^f|}{|Q^f|}, \qquad \text{if } Q^f \neq 0. \tag{12}$$

If $Q^f = 0$, the absolute discrepancy $|Q^c - Q^f|$ is used instead. This metric quantifies the extent to which a downsampling operator preserves the global integral of the field. It is particularly relevant for conserved variables such as density, momentum, or total energy.

Intra-Block Cell-Centered Error. A second family of metrics evaluates how well the coarse value $u^c(C)$ represents the distribution of fine-cell values inside the corresponding coarse control volume. For each coarse cell C, we define the blockwise root-mean-square error (RMSE) as

$$\text{RMSE}_{\text{blk}}(C) = \left(\frac{1}{N_C} \sum_{c_f \in \mathcal{F}(C)} \left(u^f(c_f) - u^c(C) \right)^2 \right)^{1/2}, \tag{13}$$

and the blockwise mean absolute error (MAE) as

$$\text{MAE}_{\text{blk}}(C) = \frac{1}{N_C} \sum_{c_f \in \mathcal{F}(C)} \left| u^f(c_f) - u^c(C) \right|. \tag{14}$$

The corresponding global indicators are obtained by averaging over all coarse cells:

$$\overline{\text{RMSE}}_{\text{blk}} = \frac{1}{|\Omega^c|} \sum_{C \in \Omega^c} \text{RMSE}_{\text{blk}}(C), \tag{15}$$

$$\overline{\text{MAE}}_{\text{blk}} = \frac{1}{|\Omega^c|} \sum_{C \in \Omega^c} \text{MAE}_{\text{blk}}(C). \tag{16}$$

These metrics measure the representativeness of the coarse cell value with respect to the unresolved fine-scale variability inside each block. The RMSE is more sensitive to large deviations, whereas the MAE provides a more robust measure of the average discrepancy.

Interface-Based Error. Because the finite-volume discretization relies on numerical fluxes through cell faces, it is also necessary to evaluate how accurately the downsampled field represents interface quantities. Let $u_{\text{int}}^{f,\alpha}$ denote the fine-grid interface value in direction α, reconstructed from the fine cell-centered field,

and let $u_{\mathrm{int}}^{c,\alpha}(C)$ denote the corresponding coarse interface value associated with coarse cell C.

For each direction α and each coarse cell C, let $\mathcal{I}_\alpha(C)$ denote the set of fine interfaces associated with the corresponding coarse interface, and let $N_{I,\alpha}(C) = |\mathcal{I}_\alpha(C)|$. We define

$$\mathrm{RMSE}_{\mathrm{face}}^{(\alpha)}(C) = \left(\frac{1}{N_{I,\alpha}(C)} \sum_{i_f \in \mathcal{I}_\alpha(C)} \left(u_{\mathrm{int}}^{f,\alpha}(i_f) - u_{\mathrm{int}}^{c,\alpha}(C) \right)^2 \right)^{1/2}, \tag{17}$$

and

$$\mathrm{MAE}_{\mathrm{face}}^{(\alpha)}(C) = \frac{1}{N_{I,\alpha}(C)} \sum_{i_f \in \mathcal{I}_\alpha(C)} \left| u_{\mathrm{int}}^{f,\alpha}(i_f) - u_{\mathrm{int}}^{c,\alpha}(C) \right|. \tag{18}$$

The global interface indicators are then obtained by averaging over all coarse cells and all coordinate directions:

$$\overline{\mathrm{RMSE}}_{\mathrm{face}} = \frac{1}{N_{\mathrm{dir}} \, |\Omega^c|} \sum_{\alpha=1}^{N_{\mathrm{dir}}} \sum_{C \in \Omega^c} \mathrm{RMSE}_{\mathrm{face}}^{(\alpha)}(C), \tag{19}$$

$$\overline{\mathrm{MAE}}_{\mathrm{face}} = \frac{1}{N_{\mathrm{dir}} \, |\Omega^c|} \sum_{\alpha=1}^{N_{\mathrm{dir}}} \sum_{C \in \Omega^c} \mathrm{MAE}_{\mathrm{face}}^{(\alpha)}(C), \tag{20}$$

where $N_{\mathrm{dir}} = 2$ in two-dimensional configurations and $N_{\mathrm{dir}} = 3$ in three-dimensional configurations.

These interface-based metrics are especially important in CFD applications because they quantify the extent to which a downsampled field preserves the quantities entering the flux evaluation. A method may perform well in terms of cell-centered averages while still introducing large errors at interfaces, which can significantly affect the numerical solution when the coarse field is used for further computation or modeling.

Interpretation of the Metrics. The conservation error measures the preservation of integral quantities and therefore reflects the physical consistency of the downsampling operator with the finite-volume formulation. The blockwise RMSE and MAE quantify how well a single coarse value represents the unresolved fine-scale variability within each coarse cell. The face-based RMSE and MAE instead evaluate the consistency of the downsampled field with the interface quantities that govern convective and diffusive fluxes. Taken together, these metrics provide a complementary assessment of conservation, representativeness, and numerical relevance.

5.2 Computational Tests

In this section, we assess the different downsampling strategies on a sequence of LES snapshots from the COVO test case. COVO (COnvection of a VOrtex)

is a basic test used to verify the correct functioning of convective schemes. In this test, a vortex is transported through a periodic grid without deformation or dissipation, and the scheme must ensure that the transport occurs at space and time scales suitable for simulating other cases of interest. The analysis is carried out on structured two-dimensional data, with fine-grid resolution 160×160 and coarse-grid resolution 40×40, corresponding to a downsampling factor equal to 4 in each spatial direction.

The variables considered are the conservative fields

$$\rho, \qquad \rho u, \qquad \rho v, \qquad \rho w, \qquad \rho E,$$

together with the temperature T. Since the underlying solver adopts a finite-volume formulation, the stored fine-grid variables are interpreted as cell-averaged quantities. A meaningful restriction operator should therefore preserve, as much as possible, both the discrete integral of conservative quantities and the local structure relevant to the evaluation of numerical fluxes.

The four downsampling operators compared in this study are:

- arithmetic averaging (`mean`),
- distance-weighted averaging (`weight`),
- minimum-based restriction (`min`),
- maximum-based restriction (`max`).

For each variable and each snapshot, we compute the following indicators:

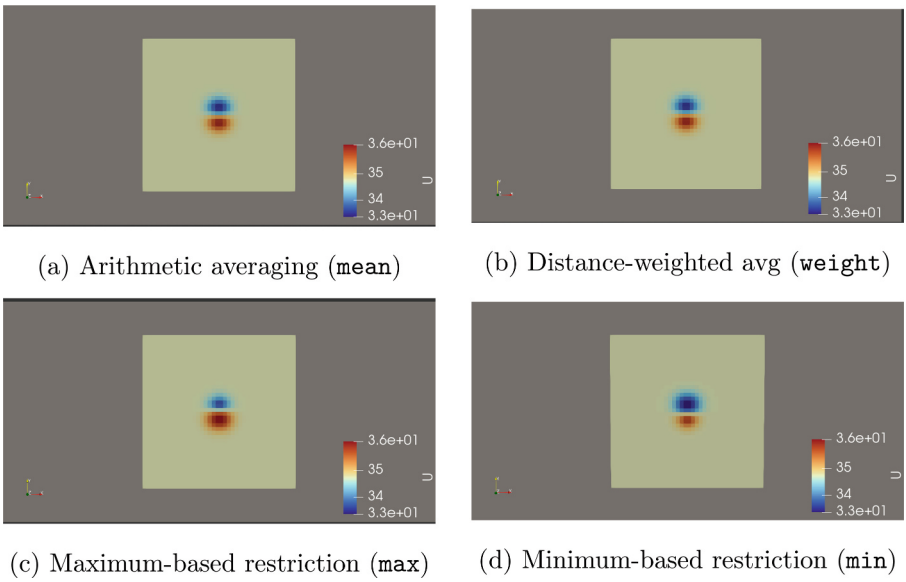
- the global conservation error $\varepsilon_{\text{cons}}$,
- the blockwise cell-centered errors $\overline{\text{RMSE}}_{\text{blk}}$ and $\overline{\text{MAE}}_{\text{blk}}$,
- the face-based errors $\overline{\text{RMSE}}_{\text{face}}$ and $\overline{\text{MAE}}_{\text{face}}$.

5.3 Qualitative Comparison of the Coarse Fields

Figures 1, 2 and 3 show a qualitative comparison of the downsampled streamwise velocity field u at the initial, intermediate and final stage of the flow evolution for the COVO test case. The final snapshot is selected as the representative visual example because it corresponds to a more evolved state of the flow, in which the relevant structures are more clearly developed and the differences among the restriction operators are more evident.

The arithmetic and distance-weighted averages produce very similar coarse fields and both preserve the overall dipolar structure of the fine solution. Nevertheless, the arithmetic average yields a slightly more balanced representation of the positive and negative lobes. By contrast, the extrema-based operators visibly distort the field: the maximum-based restriction tends to overemphasize the positive lobe, whereas the minimum-based restriction enhances the negative one. This behavior is consistent with the mathematical definition of the operators, since `max` and `min` do not compute representative coarse values but instead select local extrema within each coarse block.

Overall, the qualitative comparison already suggests that arithmetic averaging provides the most physically meaningful coarse representation, while the weighted average remains a close alternative and the extrema-based operators are less suitable for consistent state transfer across resolutions.



(a) Arithmetic averaging (`mean`) (b) Distance-weighted avg (`weight`)

(c) Maximum-based restriction (`max`) (d) Minimum-based restriction (`min`)

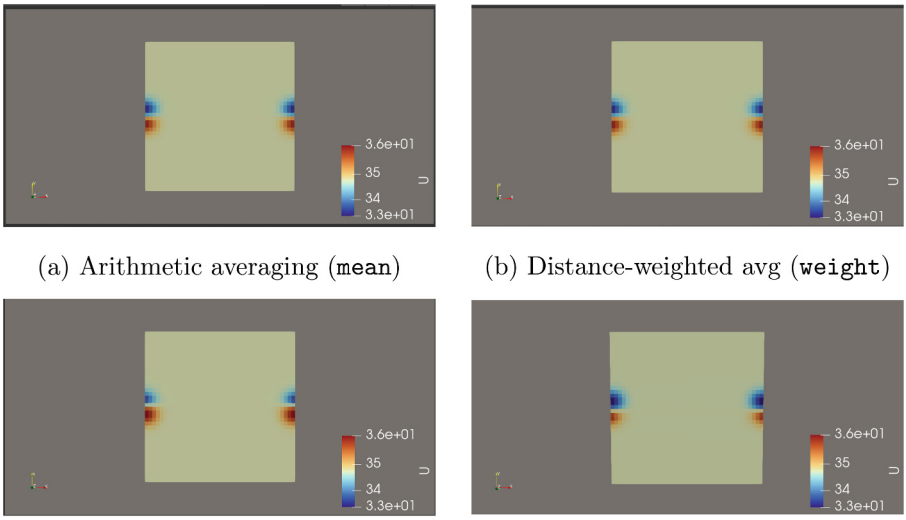
Fig. 1. Downsampling comparison at the **initial time step**.

5.4 Representative Quantitative Results

Table 1 reports the quantitative error metrics for the final snapshot, which is taken as the representative reference case. This choice is motivated by the fact that the final state corresponds to a more developed flow configuration, while still exhibiting the same ranking among the downsampling operators as observed at the initial and intermediate times.

The quantitative results consistently indicate that arithmetic averaging provides the best overall performance. For the conservative variables ρ, ρu, and ρE, the arithmetic average preserves the discrete integral up to machine precision and also yields the smallest blockwise and face-based errors. The distance-weighted strategy produces errors that are often very close to those of the arithmetic mean, but remains systematically less conservative. By contrast, the minimum- and maximum-based operators lead to significantly larger discrepancies, both in the cell-centered and in the interface-based metrics.

A particularly relevant observation concerns the face-based errors. Since numerical fluxes in the finite-volume solver are constructed from interface quantities, the quality of the face representation is directly related to the suitability of a



(a) Arithmetic averaging (`mean`) (b) Distance-weighted avg (`weight`)

(c) Maximum-based restriction (`max`) (d) Minimum-based restriction (`min`)

Fig. 2. Downsampling comparison at an **intermediate time step**.

downsampling operator for subsequent numerical use. In this respect, arithmetic averaging again emerges as the most reliable option, with the distance-weighted strategy remaining competitive but consistently slightly inferior, whereas the extrema-based operators exhibit substantially larger errors.

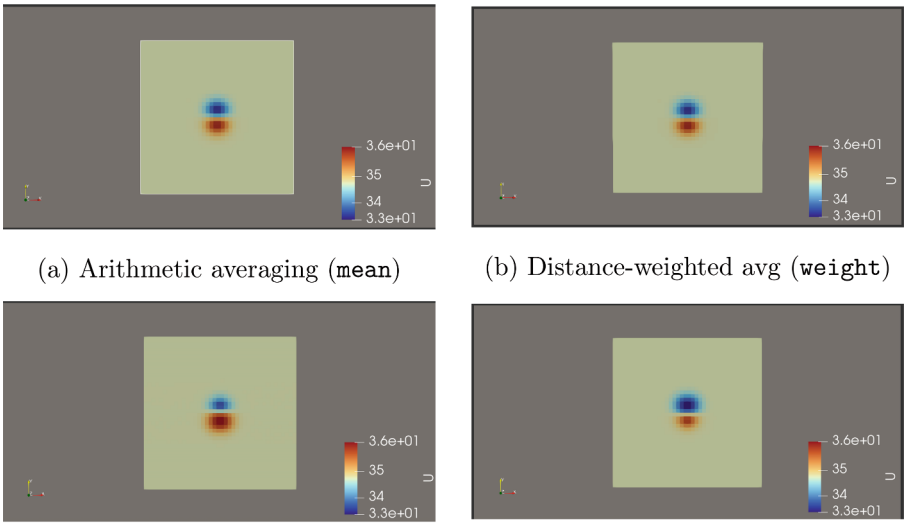
5.5 Consistency Across the Flow Evolution

The same analysis was repeated for the initial, intermediate, and final snapshots of the COVO test case. Although the absolute values of the errors vary as the flow evolves, the qualitative ranking of the downsampling operators remains unchanged. In particular, arithmetic averaging always provides the best or nearly best overall performance, distance-weighted averaging remains a close but non-conservative alternative, and the extrema-based operators systematically yield larger errors.

This temporal consistency indicates that the conclusions are not tied to a specific flow realization, but rather reflect an intrinsic property of the restriction operators themselves. The final state is therefore used in the main quantitative presentation because it offers the most developed and visually informative flow structures, while remaining fully representative of the trends observed throughout the simulation.

6 Conclusions

In this work, we investigated several downsampling strategies for LES data generated by high-fidelity finite-volume CFD solvers, with particular attention to



(a) Arithmetic averaging (`mean`) (b) Distance-weighted avg (`weight`)

(c) Maximum-based restriction (`max`) (d) Minimum-based restriction (`min`)

Fig. 3. Downsampling comparison at the **final time step**.

their suitability for cell-centered fields arising from SPARK-LES-type computations. The study was motivated by the need to transfer information from fine to coarse grids in a way that remains consistent with the numerical structure of the solver and with the physical interpretation of LES data as already filtered flow fields.

We first framed the problem in the broader context of downsampling for CFD, highlighting that restriction operators are not interchangeable: different choices preserve different properties, including conservation, boundedness, and spectral content. Starting from this perspective, we analyzed four grid-to-grid strategies on structured meshes, namely arithmetic averaging, distance-weighted averaging, minimum-based restriction, and maximum-based restriction, under the assumption that each coarse cell is the union of a set of fine cells.

The numerical results on the two-dimensional COVO test case show a clear and stable hierarchy among the considered operators. Arithmetic averaging consistently provides the best overall behavior. In particular, it preserves the integral of conservative quantities up to machine precision on uniform grids and yields the lowest blockwise and face-based errors for the variables considered. This confirms that arithmetic averaging is the most coherent choice when the objective is to obtain a coarse representation compatible with the conservative finite-volume structure of the underlying solver. Distance-weighted averaging emerges as a reasonable alternative when a more localized representation is desired. Its performance remains close to that of the arithmetic mean in many of the tested quantities, both qualitatively and quantitatively. By contrast, the extrema-based operators, i.e. minimum- and maximum-based restriction, sys-

Table 1. Representative comparison of downsampling strategies for the final flow state of the COVO test case. The columns correspond to arithmetic averaging (`mean`), distance-weighted averaging (`weight`), minimum-based restriction (`min`), and maximum-based restriction (`max`). For ρv, the relative conservation error may become poorly informative because the domain-integrated quantity is close to zero. In this case, the blockwise and face-based errors provide a more robust basis for comparison. We will use the notation $\varepsilon_{\text{cons}}(\cdot)$, $\overline{\text{RMSE}}_{\text{blk}}(\cdot)$, $\overline{\text{MAE}}_{\text{blk}}(\cdot)$, $\overline{\text{RMSE}}_{\text{face}}(\cdot)$ and $\overline{\text{MAE}}_{\text{face}}(\cdot)$, where $\cdot$ can be $\rho, \rho u, \rho v$ or ρE, to denote the corresponding error for the considered variable.

Metric(Variable)	`mean`	`weight`	`min`	`max`
$\varepsilon_{\text{cons}}(\rho)$	8.456792×10^{-16}	6.898326×10^{-14}	1.491120×10^{-7}	1.278181×10^{-7}
$\overline{\text{RMSE}}_{\text{blk}}(\rho)$	9.625817×10^{-8}	9.634005×10^{-8}	2.005217×10^{-7}	1.789093×10^{-7}
$\overline{\text{MAE}}_{\text{blk}}(\rho)$	8.179648×10^{-8}	8.176986×10^{-8}	1.753982×10^{-7}	1.503505×10^{-7}
$\overline{\text{RMSE}}_{\text{face}}(\rho)$	6.123171×10^{-8}	6.141042×10^{-8}	9.988677×10^{-8}	9.989998×10^{-8}
$\overline{\text{MAE}}_{\text{face}}(\rho)$	5.558632×10^{-8}	5.562534×10^{-8}	7.869585×10^{-8}	7.874203×10^{-8}
$\varepsilon_{\text{cons}}(\rho u)$	1.302656×10^{-14}	6.506816×10^{-10}	3.145298×10^{-4}	3.145655×10^{-4}
$\overline{\text{RMSE}}_{\text{blk}}(\rho u)$	7.669575×10^{-3}	7.673540×10^{-3}	1.499306×10^{-2}	1.499423×10^{-2}
$\overline{\text{MAE}}_{\text{blk}}(\rho u)$	6.547496×10^{-3}	6.540503×10^{-3}	1.284659×10^{-2}	1.284805×10^{-2}
$\overline{\text{RMSE}}_{\text{face}}(\rho u)$	4.795590×10^{-3}	4.803121×10^{-3}	7.996384×10^{-3}	7.996010×10^{-3}
$\overline{\text{MAE}}_{\text{face}}(\rho u)$	4.270746×10^{-3}	4.270813×10^{-3}	6.391113×10^{-3}	6.390626×10^{-3}
$\varepsilon_{\text{cons}}(\rho v)$	1.081697×10^{-3}	1.839160×10^{7}	2.815816×10^{12}	2.813619×10^{12}
$\overline{\text{RMSE}}_{\text{blk}}(\rho v)$	7.683832×10^{-3}	7.688001×10^{-3}	1.502378×10^{-2}	1.501530×10^{-2}
$\overline{\text{MAE}}_{\text{blk}}(\rho v)$	6.560065×10^{-3}	6.553028×10^{-3}	1.287379×10^{-2}	1.286374×10^{-2}
$\overline{\text{RMSE}}_{\text{face}}(\rho v)$	4.805584×10^{-3}	4.813296×10^{-3}	8.012499×10^{-3}	8.012188×10^{-3}
$\overline{\text{MAE}}_{\text{face}}(\rho v)$	4.279609×10^{-3}	4.279862×10^{-3}	6.403712×10^{-3}	6.403241×10^{-3}
$\varepsilon_{\text{cons}}(\rho E)$	2.390365×10^{-15}	5.255124×10^{-13}	1.636917×10^{-7}	1.425787×10^{-7}
$\overline{\text{RMSE}}_{\text{blk}}(\rho E)$	9.205312×10^{-3}	9.213652×10^{-3}	1.901099×10^{-2}	1.716526×10^{-2}
$\overline{\text{MAE}}_{\text{blk}}(\rho E)$	7.827272×10^{-3}	7.825163×10^{-3}	1.658196×10^{-2}	1.444322×10^{-2}
$\overline{\text{RMSE}}_{\text{face}}(\rho E)$	5.832684×10^{-3}	5.851590×10^{-3}	9.459973×10^{-3}	9.462631×10^{-3}
$\overline{\text{MAE}}_{\text{face}}(\rho E)$	5.323460×10^{-3}	5.328789×10^{-3}	7.425071×10^{-3}	7.431006×10^{-3}

tematically produce larger discrepancies. Although these operators may still be useful for feature-detection purposes or for bounding analyses, they do not provide representative coarse values in the finite-volume sense and significantly distort both the cell-centered and interface-based description of the flow. This makes them less appropriate for consistent state transfer across resolutions or for subsequent use inside a numerical solver. An important outcome of the study is that the ranking of the operators remains unchanged throughout the flow evolution. The same qualitative and quantitative trends are observed at the initial, intermediate, and final stages of the simulation, indicating that the conclusions are not tied to a particular snapshot but rather reflect intrinsic properties of the restriction procedures themselves. This temporal consistency strengthens the

robustness of the analysis and supports the use of arithmetic averaging as the reference downsampling strategy for this class of data.

The present work constitutes a first step toward the definition of reliable restriction operators for LES-based CFD workflows and, more generally, for future integration with data-driven models acting on coarse–grid representations. The proposed restriction operators are defined in a dimension-independent manner, as they rely on the aggregation of cell-averaged quantities over control volumes, and are therefore directly applicable to fully three–dimensional structured grids. As a next step, we are extending the analysis to the three-dimensional Taylor–Green vortex (see [13]), which represents a more demanding benchmark due to its fully three-dimensional vortex dynamics and richer inter–scale interactions. Preliminary results indicate that the relative performance of the considered downsampling operators remains consistent with the two-dimensional case, with arithmetic and distance-weighted averaging outperforming the extrema-based strategies. For brevity, the corresponding three-dimensional results are not reported here, but we underline that they further support the robustness of the observed trends. Although the study is partly motivated by data-driven CFD applications, its primary objective is to identify downsampling strategies that preserve the physical and numerical consistency of the underlying finite-volume formulation. This aspect is particularly relevant for machine–learning–based approaches, where the quality and physical coherence of the training data play a crucial role in ensuring the reliability of the resulting models.

Acknowledgments. We thank Professor Francesco Capuano (Polytechnic University of Catalonia), Professor Raffaele D'Ambrosio (University of L'Aquila) and Professor Anna Esposito (University of Campania) for the fruitful discussions and useful comments. We also thank the reviewers for their valuable comments. The research of Gerardo Toraldo is partially supported by INdAM-GNCS and the research of Ferdinando Zullo is partially supported by INdAM-GNSAGA.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Berkooz, G., Holmes, P., Lumley, J.L.: The proper orthogonal decomposition in the analysis of turbulent flows. Annu. Rev. Fluid Mech. **25**, 539–575 (1993)
2. Capuano, F., Catalano, P., Mastellone, A.: Large-eddy simulation of a spatially-evolving turbulent mixing layer, APS Division of Fluid Dynamics Meeting Abstracts, R03-005 (2015)
3. Capuano, F., Mastellone, A., De Angelis, E.M.: A conservative overlap method for multi-block parallelization of compact finite-volume schemes. Comput. Fluids **159**, 327 (2017)
4. Capuano, F., Mastellone, A., Benedetto, S.D., Cutrone, L., Schettino, A.: Preliminary developments towards a high-order and efficient LES code for propulsion applications. In: Proceedings of 6th European Conference on Computational Fluid Dynamics (ECFD VI), Barcelona, Spain (2014)

5. Deconinck, H., Koren, B.: Algorithm development in CFD: Multidimensional upwinding and multiple semi-coarsening multigrid. In: Computational methods in applied sciences' 96, Paris, 9-13 September 1996 (1996)
6. Fukami, K., Fukagata, K., Taira, K.: Super-resolution reconstruction of turbulent flows with machine learning. J. Fluid Mech. **870**, 106–120 (2019)
7. Germano, M.: Turbulence: the filtering approach. J. Fluid Mech. **238**, 325–336 (1992)
8. Gullbrand, J., Chow, F.K.: The effect of numerical errors and turbulence models in large-eddy simulations of channel flow. J. Fluid Mech. **495**, 323–341 (2003)
9. Harten, A.: Multiresolution algorithms for the numerical solution of hyperbolic conservation laws. Commun. Pure Appl. Math. **48**(12), 1305–1342 (1995)
10. Kim, H., et al.: Unsupervised deep learning for super-resolution reconstruction of turbulence. J. Fluid Mech. **910**, A29 (2021)
11. Pope, S.B.: Turbulent flows. Meas. Sci. Technol. **12**(11), 2020–2021 (2001)
12. Kochkov, D., et al.: Machine learning–accelerated computational fluid dynamics. Proceedings of the National Academy of Sciences, vol. 118, no. 21 (2021)
13. Mastellone, A., Capuano, F., Benedetto, S.D., Cutrone, L.: Direct numerical simulation of the Taylor-Green Vortex at Re = 1600. In: Proceedings of 3rd International Workshop on High-Order CFD Methods, Kissimmee, USA (2015)
14. Pulido, A., et al.: Survey and analysis of multiresolution methods for turbulence data. Comput. Fluids **125**, 39–58 (2016)
15. Radhakrishnan, S., Bellan, J.: Explicit filtering to obtain grid-spacing-independent and discretization-order-independent large-eddy simulation of compressible single-phase flow. J. Fluid Mech. **697**, 399–435 (2012)
16. Sagaut, P.: Large Eddy Simulation for Incompressible Flows: an introduction. Springer, Heidelberg (2006)
17. Stüben, K.: A review of algebraic multigrid. In: Numerical Analysis: Historical Developments in the 20th Century, pp. 331–359 (2001)
18. Unat, D., Hromadka, T., Baden, S.B.: An adaptive sub-sampling method for in-memory compression of scientific data. In: Data Compression Conference. IEEE (2009)

Numerical Methods and Machine Learning in Science & Engineering (NMLSE 2026)

A Comparative Study of Optimization Models for Superellipsoid Fitting in 3D Object

Maria Silva[1], Fernanda Costa[1,2], and Eliana Costa e Silva[3,4(✉)]

[1] Department of Mathematics, University of Minho, Braga, Portugal
[2] Center of Mathematics (CMAT), University of Minho, Braga, Portugal
[3] CIICESI, ESTG, Instituto Politécnico do Porto, Felgueiras, Portugal
eos@estg.ipp.pt
[4] Algoritmi Center, University of Minho, Guimarães, Portugal

Abstract. This work presents an in-depth analysis of superquadrics and their application in the modeling of three-dimensional objects. The study focuses on object modeling from point cloud data, a key problem in robotics and computer vision. Two optimization models are proposed, based on implicit surface fitting and volume minimization, respectively. Both formulations are solved using the SLSQP method. Results are presented through graphical visualizations, correlation analysis for different upper limits of the upper bound of the parameters of the superquadrics, followed by a comparative analysis using volume-based metrics to assess the quality of fit. The results show that, for the studied objects, the best results are obtained for the volume-based model. Further, the influence of upper limits is significant for the implicit surface fitting model but not for the volume-based. These findings reinforce that the volume-based formulation offers a more robust and reliable approach for superellipsoid fitting across the tested objects.

Keywords: Superellipsoids · Optimization · SLSQP · 3D Modeling

1 Introduction

Three-dimensional (3D) object representation is a fundamental problem in several scientific and technological domains, including robotics, computer vision, and computer graphics. The ability to accurately model object geometry is essential for enabling autonomous systems to interact effectively with the physical world. In robotics, this capability is particularly relevant for tasks such as object manipulation, grasp planning, navigation, and collision avoidance, where the quality of geometric representations directly impacts performance and safety.

Robotic systems typically rely on sensor data, such as point clouds acquired from depth cameras or LiDAR, to perceive their environment. Converting this raw data into compact and meaningful geometric models remains a challenging task. On one hand, dense representations such as meshes or raw point clouds

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 241–254, 2027.
https://doi.org/10.1007/978-3-032-30530-5_15

provide high fidelity but are computationally expensive. On the other hand, simpler geometric primitives offer efficiency but often lack the flexibility required to accurately describe complex shapes. Therefore, there is a need for representations that balance expressiveness and computational tractability.

Superquadrics have emerged as a promising solution to this problem. As a family of parametric surfaces, superquadrics generalize classical quadrics by introducing shape parameters that control curvature and surface sharpness. This allows them to represent a wide range of geometries—from smooth, rounded objects to more angular and box-like shapes—using a relatively small number of parameters. In addition to shape, superquadrics can encode position, orientation, and scale, making them suitable for modeling objects in arbitrary configurations within a 3D space.

Their compact representation and analytical formulation make superquadrics particularly attractive for robotics applications. Tasks such as collision detection, distance computation, and grasp planning can benefit from the mathematical properties of superquadrics, which enable efficient geometric reasoning. Moreover, their parametric nature facilitates optimization-based fitting to sensor data, allowing objects to be reconstructed directly from point clouds.

Despite these advantages, most existing approaches have focused primarily on superellipsoids, a specific subclass of superquadrics. While effective for modeling convex objects, superellipsoids are limited in their ability to represent more complex geometries, such as objects with holes or non-convex features. This limitation motivates the exploration of additional superquadric forms, including supertoroids and superhyperboloids, which provide greater flexibility in shape representation.

In this work, we aim to extend the use of superquadrics for 3D object modeling by exploring a broader family of shapes and developing optimization-based methods for fitting these models to point cloud data. The proposed approach seeks to improve both the accuracy and generality of superquadric-based representations, enabling their application in more complex and realistic scenarios.

2 Related Work

The problem of 3D object representation has been widely studied, with approaches ranging from explicit geometric models to data-driven methods. Traditional representations include polygonal meshes and voxel grids, which provide detailed descriptions of object geometry but often incur high computational costs. More compact representations based on geometric primitives, such as spheres, cylinders, and ellipsoids, have been proposed to address this issue, although they typically lack the flexibility required for complex shapes.

Superquadrics were introduced as a generalization of quadric surfaces, providing a unified framework for representing a wide variety of shapes using a small set of parameters [2]. Early work demonstrated their effectiveness in computer graphics and vision, particularly for object recognition and modeling tasks. Their ability to smoothly interpolate between different geometric forms makes them a versatile tool for shape approximation.

In robotics, superquadrics have been applied to object modeling and scene understanding using point cloud data. For instance, Costa e Silva et al. [8] demonstrated the use of superquadrics to model objects in robotic workspaces, showing improvements over simpler representations such as ellipsoids in terms of both accuracy and compactness. Their approach highlighted the potential of superquadrics for tasks such as grasp planning and collision detection. However, their study was limited to superellipsoids, restricting the range of shapes that could be effectively modeled.

Subsequent research has explored optimization techniques for fitting superquadrics to data. These methods typically formulate the fitting problem as a non-linear optimization task, where the objective is to minimize the distance between the model surface and the observed data points [7]. Common approaches include least-squares formulations and variants based on implicit surface representations. While effective, these methods can be sensitive to initialization and may struggle with noisy or incomplete data.

Other works have investigated alternative objective functions and constraints to improve robustness. For example, volume-based metrics have been proposed to ensure that the fitted superquadric provides a tight approximation of the object while avoiding excessive overestimation [10]. These approaches aim to balance accuracy with stability in the optimization process.

Despite these advances, the literature still lacks comprehensive studies that consider a wider range of superquadric families, such as supertoroids and superhyperboloids, in a unified framework. These shapes offer additional modeling capabilities, particularly for objects with non-convex geometries or complex topologies. Furthermore, there is a need for methodologies that can generalize across different object types and operate effectively under real-world conditions, including noise, occlusions, and varying sensor quality.

The present work addresses these gaps by systematically exploring multiple superquadric forms and comparing different optimization strategies for fitting them to point cloud data. By extending beyond superellipsoids and incorporating more expressive shape families, this study aims to advance the state of the art in superquadric-based object modeling and contribute to the development of more capable robotic perception systems.

Superquadrics stand out in the literature as a versatile and efficient mathematical approach for representing a wide range of three-dimensional geometric shapes [5,7,8]. Due to their closed-form equations, they provide a compact and computationally efficient description, which makes them particularly relevant in areas such as robotics, computer vision, and object modeling.

Studies such as those by [8] and [7] demonstrate the applicability of variants such as superellipsoids, super hyperboloids, and supertoroids, highlighting the advantages of implicit parameterization, which facilitates both modeling and subsequent geometric analysis. In turn, in [5] the authors propose a generalization to hyperquadrics, increasing the representation capability of shapes, albeit with greater mathematical complexity.

However, traditional superquadrics present limitations, especially when applied to modeling objects with irregular contours or non-convex geometries. They are particularly well-suited to representing symmetric or regular-shaped objects but struggle to handle structures that contain cavities or holes [5,7,8].

In order to overcome these limitations, [11] proposes the use of supertoroids—a fourth-degree generalization of superquadrics—capable of modeling objects with a single hole. Their approach allows for a straightforward derivation of the intrinsic and extrinsic parameters of the surface, based on direct expressions for the major and minor radii, while maintaining a computationally efficient distance function. This advance significantly expands the range of shapes that can be geometrically modeled, showing promise for applications in robotic manipulation and complex scene generation.

Additionally, [6] demonstrates the practical applicability of superquadrics in real-world scenarios, proposing a fast and effective method for grasping unknown objects in unstructured environments. The approach is based on partial reconstruction of objects from incomplete views, assuming symmetry to generate approximate models to which superquadrics are fitted. Using the obtained geometric parameters, the system determines the contact points and the optimal approach direction for an optimized antipodal grasp. Tests on PR2 (Personal Robots 2) robots showed that this solution, despite using simplified representations, outperforms algorithms based on neural networks in terms of temporal efficiency and accuracy in cluttered environments.

In parallel, the automation of the superquadric fitting process has been explored using deep learning techniques. [9] presents an approach based on convolutional neural networks, operating in two stages: in a first step, a Mask R-CNN (Region-based Convolutional Neural Network) architecture identifies regions with structures compatible with superquadrics in depth scenes; in a second step, a CNN regressor (Convolutional Neural Network) fits the parametric models to the segmented regions. This method has proven effective on both synthetic and real data, enabling the description of complex shapes with a reduced set of interpretable parameters and with significantly lower processing time compared to conventional techniques.

This evolution, from classical formulations to practical applications and automated neural network-based approaches, highlights the growing importance of superquadrics in modern robotics, both for perception and representation tasks, as well as for physical actions such as grasping.

3 Methodology

This section presents the methodology adopted for modeling three-dimensional objects using superquadrics. Motivated by the need for accurate and efficient object representations in robotic applications, the proposed approach focuses on fitting superquadric surfaces, more specifically superellipsoids, to point clouds with high accuracy, a key requirement for perception, collision detection, and object manipulation.

For modeling three-dimensional objects using superquadrics two nonlinear optimization models are considered. Model 1, based on the approach proposed in [8], employs the inside–outside function as the objective. Model 2, estimates the parameters of the superellipsoid using a volume-based objective function.

The optimization method used to solve both models is presented in Subsect. 3.3.

3.1 Model 1

In the first model, the objective function is based on the implicit formulation of the superquadric, inspired by the approach proposed in [8]. Specifically, the objective function is defined as the mean fitting error of each point relative to the surface, multiplied by a volume based normalization term. This normalization prevents the objective function from favoring solutions with disproportionate scales, effectively introducing a regularization term proportional to the approximate volume of the fitted superquadric. The imposed constraints ensure that all N data points lie inside or on the superquadric surface.

The exact formulation of the optimization problem is:

$$\min_{\lambda \in \mathbb{R}^5} \; F(\boldsymbol{\lambda}) = \frac{\sqrt{\lambda_0 \lambda_1 \lambda_2}}{N} \sum_{i=1}^{N} \left[f^{\frac{\lambda_3}{2}}(\boldsymbol{\lambda}, \mathbf{C}_i) - 1 \right]^2 \tag{1}$$

$$\text{s.t.} \;\; f(\boldsymbol{\lambda}, \mathbf{C}_i) \leq 1, \quad \forall i = 1, \ldots, N. \tag{2}$$

$$\boldsymbol{\lambda}^{(L)} \leq \boldsymbol{\lambda} \leq \boldsymbol{\lambda}^{(U)}, \tag{3}$$

with

$$f(\boldsymbol{\lambda}, \mathbf{C}_i) = \left(\left(\frac{x}{\lambda_0} \right)^{\frac{2}{\lambda_4}} + \left(\frac{y}{\lambda_1} \right)^{\frac{2}{\lambda_4}} \right)^{\frac{\lambda_4}{\lambda_3}} + \left(\frac{z}{\lambda_2} \right)^{\frac{2}{\lambda_3}},$$

where $\boldsymbol{\lambda}$ is the parameter vector of the superellipsoid. Specifically, λ_0, λ_1, and λ_2 are the scale parameters along the x, y, and z axes, respectively, λ_3 and λ_4 are the shape parameters that control the "roundness" or "squareness" of the surface.

3.2 Model 2

The second model, aims at fitting a superquadric that encloses the point cloud, which minimizes the volume of the enclosing superquadric. Therefore, the constraints are the same as in Model 1 but the objective function is given as the volume of the superquadric. Specifically, the objective function is:

$$F(\boldsymbol{\lambda}) = 2\lambda_0 \lambda_1 \lambda_2 \lambda_3 \lambda_4 \, B\left(\tfrac{\lambda_3}{2} + 1, \lambda_3\right) \, B\left(\tfrac{\lambda_4}{2}, \tfrac{\lambda_4}{2}\right), \tag{4}$$

where $B(x, y)$ is the beta function defined as:

$$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1} \, dt.$$

This formulation makes it possible to find the superellipsoid with the smallest possible volume that still satisfies the condition of containing all the points, resulting in more compact solutions with restricted volume, which can be particularly useful in applications that require minimal boundaries or spatial optimization

3.3 The SLSQP Method

The Sequential Least Squares Quadratic Programming (SLSQP) algorithm was originally developed by Dieter Kraft in the 1980s [4], based on the sequential quadratic programming (SQP) method. SLSQP is particularly well suited for nonlinear constrained optimization problems, provided that both the objective and constraint functions are sufficiently smooth (i.e., continuously differentiable), as established in Kraft's original formulation [4]. The current Python implementation - PySLSQP- was designed to handle medium to large-scale problems and has been successfully applied in complex industrial contexts [3].

The implementation used in this work is based on the PySLSQP package developed by Joshy and Hwang [3]. In this version, subproblems are solved using quadratic approximations with linearized constraints, and the Hessian matrix is updated through quasi-Newton approximations, specifically the BFGS (Broyden–Fletcher–Goldfarb–Shanno) algorithm, as described in the implementation's documentation [3]. At each iteration, SLSQP performs the following main steps:

1. computing the *gradients* of the objective and the constraint functions;
2. building a local quadratic approximation of the objective function;
3. linearization of the nonlinear constraints;
4. solving a quadratic programming subproblem with linear constraints;
5. updating the solution based on the subproblem results.

Convergence of the method is guaranteed under continuity and differentiability of the functions involved. The method may also employ quasi-Newton approximations to avoid the explicit computation of the Hessian of the objective function. The selection of SLSQP as the optimizer was motivated by: **(a)** its ability to naturally handle nonlinear constraints, both equality and inequality; **(b)** its support for simple bound constraints on variables, which simplifies modeling; **(c)** its numerical stability and robust convergence behavior; **(d)** its availability in optimization libraries such as `SciPy` and other widely used scientific tools.

4 Results

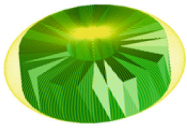
In this section, the numerical results are presented and discussed. The goal is to assess the performance of two models in fitting superquadrics to various 3D objects represented as point clouds. The objects used in the experiments are first described, followed by the experimental design, optimization method, and evaluation criteria. Results for Model 1 and Model 2 are reported separately,

including final parameter values and their influence on the objective function using Pearson correlation. A comparative analysis based on the estimated volumes of the fitted superquadrics identifies which model provides more accurate and efficient representations of the studied objects.

4.1 Experimental Setup

In the numerical experiments, three objects were used in robotic experiments, namely wheel, nut, and column, previously considered in [1]. For each object, a set of surface points is known. The characteristics of each object, specifically, the diameters along the x, y and z axes, corresponding to the maximum distance between points in the Ox, Oy and Oz planes, respectively, as well as the number of known points, are presented in Table 1.

Table 1. Characteristics of the objects used in the numerical experiments.

Parameter	nut	wheel	column
			
Diameter x	100	140	80
Diameter y	86.6	140	80
Diameter z	40	40	760
Number of Points	28	362	54

It is possible to observe that the number of available points varies significantly. Some objects have a more detailed geometric description, while others are represented more sparsely. In fact, while nut and wheel contains only 40 points, the Column contains 362. Also, the wheel and nut are flatter objects, as their dimensions along the x and y axes are much smaller than along z axis, whereas the Column is elongated, with a z dimension far greater than the remaining axes. Table 1 also shows that the objects differ greatly in scale. For instance, while the maximum dimension of the Column is 760 mm, the maximum dimension of the nut is 100 mm. Each object exhibits unique characteristics, which justify their use in the numerical experiments on superquadric fitting.

To standardize object dimensions and enable more meaningful comparisons of results across different objects, normalization was applied. Specifically, the 3D points were first separated into their respective X, Y and Z coordinates, and each coordinate was normalized independently to $[-1, 1]$, as:

$$x_i = \frac{X_i}{\max(|X|)}, \quad y_i = \frac{Y_i}{\max(|Y|)}, \quad z_i = \frac{Z_i}{\max(|Z|)},$$

where X_i, Y_i and Z_i denote the original coordinates, and x_i, y_i and z_i are the normalized ones. The dimensions of the cloud point along each axis, R_x, R_y and R_z, were computed as the difference between the maximum and minimum values of each coordinate. This procedure also ensured that all objects were centered at $(0, 0, 0)$ with orientation $(0, 0, 0)$.

Given the highly nonlinear nature of the optimization problems presented in Sect. 3, the existence of multiple optima is expected. Thus, considering that SLSQP is a local solver, preliminary tests were carried out showing that the optimal solution obtained for the superquadrics parameters depended on the bounds defined for the decision variables of the optimization problems (i.e. the parameters of the superquadrics $\boldsymbol{\lambda}$).

Varying the lower bounds of λ_3 and λ_4 did not lead to significant changes in the results. Therefore, in line with previous studies, the lower and upper limits of λ_3 and λ_4, were set fix to 0.1 and 1.4, respectively.

On the other hand, changes in the upper bounds of λ_0, λ_1, and λ_2 affected of the results, justifying their detailed analysis. Based on this, it was defined to fix the lower limits of λ_0, λ_1, λ_2 to $0.5R_x$, $0.5R_y$ e $0.5R_z$, respectively, while the upper limits would vary from $0.6R_x$ to $0.8R_x$, $0.6R_y$ to $0.8R_y$, and $0.6R_z$ to $0.8R_z$, respectively. Specifically, a total of 27 combinations for the upper bounds of λ_0, λ_1 and λ_2 were considered (see Table 2).

Table 2. Lower and upper limits for $\boldsymbol{\lambda}$.

Var.	λ^L	λ^U
λ_0	$0.5R_x$	$\{0.6R_x, 0.7R_x, 0.8R_x\}$
λ_1	$0.5\,R_y$	$\{0.6R_x, 0.7R_x, 0.8R_x\}$
λ_2	$0.5\,R_z$	$\{0.6R_x, 0.7R_x, 0.8R_x\}$
λ_3	0.1	1.4
λ_4	0.1	1.4

4.2 Results Using Model 1

For the all the objects and for the 27 combinations of the bounds of $\boldsymbol{\lambda}$ (see Table 2), an optimal solution was obtained for fitting the best superellipsoid. To improve clarity, in Table 3 and Fig. 1, only the results corresponding to the combinations of $\boldsymbol{\lambda}$ bounds that yield the smallest optimal value (referred to as the best case) and the largest optimal value (the worst case) are presented. The results show that, for all objects except the Column, the optimal objective values are close to zero ($<5 \times 10^{-6}$). This indicates that, for these objects, the clouds' point lie very close to the surface of the fitted superquadric. In contrast, for the Column, some points lie inside the superellipsoid, reflecting the greater geometric complexity of this object.

Table 3. Results of $\boldsymbol{\lambda}$ for the worst (w) and best (b) optimal objective function value F^* using Model 1 for superellipsoid fitting.

Object	Case	Optimal values					Objective	Upper limits		
		λ_0^*	λ_1^*	λ_2^*	λ_3^*	λ_4^*	F^*	$\lambda_0^{(U)}$	$\lambda_1^{(U)}$	$\lambda_2^{(U)}$
column	w	40.1872	43.640	383.677	0.1	1.1143	4.19659e-3	0.7 R_x	0.8 R_y	0.7 R_z
	b	40.1869	43.619	383.717	0.1	1.1139	4.19655e-3	0.7R_x	0.7R_y	0.6R_z
nut	w	**70.00**	**51.96**	20.048	0.114	**0.100**	4e-7	0.7 R_x	0.6 R_y	0.6 R_z
	b	**78.84**	**69.28**	20.000	0.100	**0.604**	4.72e-11	0.8 R_x	0.8 R_y	0.7R_z
wheel	w	**83.15**	**83.997**	20.03	0.100	0.740	2.189e-6	0.7 R_x	0.6 R_y	0.7 R_z
	b	**110.34**	**110.30**	20.00	0.100	0.761	3.270e-11	0.8R_x	0.8R_y	0.7R_z

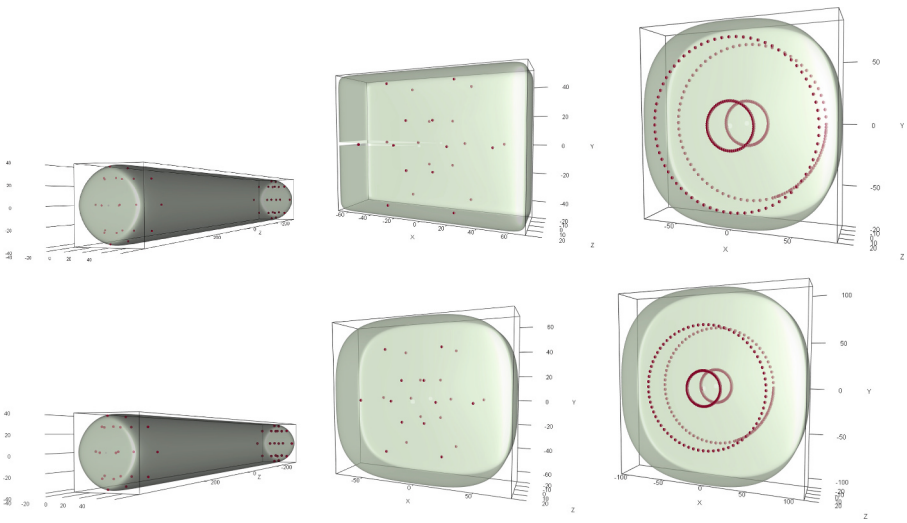


Fig. 1. Comparison between the worst (top) and best (bottom) objective-function results for the column (left), nut (middle) and wheel (right), using Model 1.

For the Column, the solutions obtained in the worst and best cases are very similar. The difference in the objective function between the two cases is only 4.8×10^{-8}. The parameters, λ_0, λ_1 and λ_2 coincide up to the second decimal place, while λ_3 is the same and λ_4 varies by 4×10^{-4}.

For the nut, the worst and best cases exhibit similar fitting behavior, with a difference in the objective function of 4×10^{-7}. The optimal values of λ_0, λ_1 and λ_2 differ by 8.84, 17.32 and 0.048, respectively, while the differences for λ_3 and λ_4 are 0.014 and 0.504.

For the wheel, both the worst and best cases result in fittings that closely match the real geometry of the object. The difference in the objective function between the two cases is 2.19×10^{-6}. In terms of parameters, λ_0 and λ_1 differ

by 27.19 and 26.30, respectively, while λ_2 varies by only 0.03. The values of λ_3 do not differ and the values of λ_4 differ by 0.021.

The results in Table 3 suggest that there is an association between the upper limits that is set for λ_0, λ_1 and λ_2 and the optimal solution. To evaluate this, Spearman correlations were computed. The Spearman correlation coefficient was selected for the association analysis because it is appropriate for data that do not follow a normal distribution. This coefficient does not require parametric assumptions and has the advantage of capturing both linear and nonlinear monotonic relationships between variables, thus providing more robust results. Table 4 presents the correlation coefficients and their corresponding p-values.

Table 4. Spearman correlation between the upper bounds of the optimization model parameters and the optimal solution value.

Object	Correlation Coefficient (p-value) of F^* *vs.*					
	$\lambda_0^{(U)}$		$\lambda_1^{(U)}$		$\lambda_2^{(U)}$	
column	$\simeq 0$	(0.99)	$\simeq 0$	(0.25)	$\simeq 0$	(0.99)
nut	-0.27	(0.17)	$\mathbf{-0.42}$	(0.03)	-0.14	(0.47)
wheel	$\mathbf{-0.53}$	(<0.01)	$\mathbf{-0.53}$	(<0.01)	0.01	(0.97)

For a significant level of 10%, no statistically significant correlation exists between the choice of the upper limit of λ_2 and the optimal solution found, as all the p-values are greater than 10%.

For the wheel, both λ_0 and λ_1 upper limits are significantly negatively associated ($r \simeq -0.53$) to the optimal solution. Thus, as the upper limit of the λ_0 and λ_1 parameters increases, the optimal solution decreases, which indicates a tighter fitting of the superellipsoid surface to the cloud points.

For the nut, the only significant correlation is between the upper limit of the λ_1 parameter of the superellipsoid, which presents a negative correlation coefficient of -0.42. This indicates that as the upper limit of λ_1 increases, the optimal objective value decreases, thus a tighter fitting of the superellipsoid surface to the cloud points is found.

These results highlight the sensitivity of the optimal solution to the upper bounds imposed on the superellipsoid parameters, but only for certain objects. This reinforces the importance of conducting an object-specific analysis to ensure an optimal fit in each case.

4.3 Model 2

Model 2 uses the volume of the superquadric as the objective function. The best and worst results of the 27 combinations of lower and upper bounds of the parameters of each object are presented Table 5 and Fig. 2.

For the Column, the difference of the solutions obtained in the worst and best cases is of 2×10^{-7}. The parameters λ_0, λ_1 and λ_2 present differences in the

Table 5. Results of $\boldsymbol{\lambda}$ for the worst (w) and best (b) optimal objective function value F^* using Model 2 for superellipsoid fitting.

Object	Case	Optimal values					Objective	Upper limits		
		λ_0^*	λ_1^*	λ_2^*	λ_3^*	λ_4^*	F^*	$\lambda_0^{(U)}$	$\lambda_1^{(U)}$	$\lambda_2^{(U)}$
column	w	40.1486153	42.224166	387.94293	0.1	1.071	6.5607275	$0.8\,R_x$	$0.6\,R_y$	$0.8\,R_z$
	b	40.1486163	42.2241964	387.94278	0.1	1.071	6.5607273	$0.8R_x$	$0.8R_y$	$0.7R_z$
nut	w	51.0240076	51.9372869	21.129346	0.1	1.046	7.9209705	$0.6\,R_x$	$0.6\,R_y$	$0.7\,R_z$
	b	51.0240077	51.936918	21.1293461	0.1	1.046	7.9209704	$0.8\,R_x$	$0.8\,R_y$	$0.8R_z$
wheel	w	71.4336643	71.4336637	21.1293462	0.1	1	6.8615689	$0.7\,R_x$	$0.6\,R_y$	$0.7\,R_z$
	b	71.4336672	71.4336641	21.1293413	0.1	1	6.8615685	$0.7\,R_x$	$0.6\,R_y$	$0.8\,R_z$

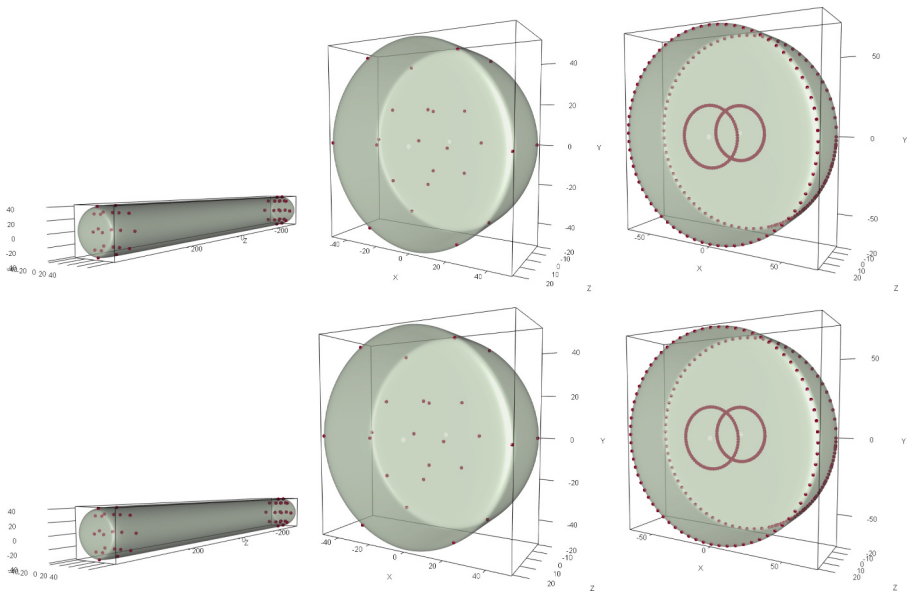


Fig. 2. Comparison between the worst (top) and best (bottom) objective function results for the column (left), nut (middle) and wheel (right), using Model 2.

order of 10^{-5}. Similar to the Column, the results for the worst and best cases for nut are almost the same. In fact, the difference in the objective function values is only 6.7×10^{-10}. In terms of the λ_0, λ_1 and λ_2 parameters, the differences occur only on the 7th decimal place. Finally, for the wheel, the difference is of 3.6×10^{-7}. The parameters λ_0, λ_1 and λ_2 present differences in the order of 10^{-6}. For all objects, λ_3 and λ_4 are the same for the best and worst cases.

These results suggest that there is no association between the upper limits that are set for λ_0, λ_1 and λ_2 and the optimal solution found. In fact, for a level of significance of 10%, the only statistically significant correlation was found

between the upper bound of λ_1 and the optimal value of the objective function, for the column. This was a moderate negative correlation of -0.32 ($p - value \simeq 0.099$). These results are not in line with the results obtained in Model 1, for which no significant correlations were found for this object.

4.4 Comparison of the Results for Models 1 and 2

Overall, the analysis of the correlations between the upper bounds of the parameters λ_0, λ_1, and λ_2 and the objective function value revealed differences in how the two models behave with respect to parameter sensitivity across the various objects. In Model 1, only a subset of the parameters significantly influenced the quality of the fit for certain objects. In Model 2, the impact of changing the upper bounds of the parameters was almost none.

To determine which of the two models provides a better representation of the objects, the estimated volume for each fitted superquadric was computed. Since all points of each object are contained within the corresponding superquadric, the resulting volume serves as an appropriate comparative indicator of fitting quality. In fact, the smaller the estimated volume for the same set of points, the better the superquadric fits the object when comparing the two models quantitatively.

Table 6 presents the volume of each superquadric obtained from the optimal fit for both models applied to the objects. Overall, it is observed that Model 2 produced significantly smaller volumes.

Table 6. Volume of the superquadric fitting each object for the worst (w) cases and best (b) for the two proposed models.

Object	Case	Model 1	Model 2	diff.
column	w	4010806.876	3988922.314	−2.19e+04
	b	4009976.797	3988922.219	−2.11e+04
	diff.	−8.30e+02	−9.45e−02	
nut	w	575621.197	342988.321	−2.33e+05
	b	781018.363	342988.317	−4.38e+05
	diff.	2.05e+05	−3.78e+05	
wheel	w	958297.984	672433.751	−2.86e+05
	b	1656186.185	672433.715	−9.84e+05
	diff.	6.98e+05	−3.57e−02	

For the column, the best solution obtained using model 1 presents a smaller volume that the one obtained for the worst case (it presents a difference of -8.30×10^2). For the other objects, the volume is larger for the best case. This can be explained by the fact that the nut and wheel have a hole. For model 2, as expected, the volume is always better for the best case.

Comparing the volumes using model 1 and 2, for all objects, as expected, the volume of the superquadric fitting the object is smaller using Model 2 than using Model 1(see last column in Table 6).

In summary, the results show that Model 2 performs better for all objects, as it produces more accurately fitted superquadrics with smaller volumes.

5 Conclusions and Future Work

This work provided a detailed examination of superellipsoids and their suitability for modeling three-dimensional objects from point cloud data. The study focuses on object modeling from point cloud data, using three different objects, previously used in human-robot collaboration tasks.

Two optimization models were developed—one based on implicit surface fitting and the other on volume minimization—and evaluated through graphical inspection, correlation analysis under varying upper-bound constraints, and a volume-based comparison of the resulting fits.

The results consistently showed that the volume-based formulation yields more compact and accurate representations, while also demonstrating reduced sensitivity to parameter upper limits when compared to the implicit surface fitting model. Overall, the findings indicate that the volume-minimization approach offers a more robust and reliable framework for superellipsoid fitting across the studied objects.

As future work, it would be valuable to extend this analysis to other families of superquadrics, enabling a broader assessment of their modeling capabilities and potential advantages in different geometric contexts.

Acknowledgments. E. Costa e Silva was supported by national funds through FCT - Fundação para a Ciência e Tecnologia through project UIDB/04728/2025 (https://doi.org/10.54499/UID/04728/2025).

References

1. Costa e Silva, E., et al.: Towards human-like bimanual movements in anthropomorphic robots: a nonlinear optimization approach (2015). https://doi.org/10.12785/amis/090210
2. Jaklic, A., Leonardis, A., Solina, F.: Segmentation and Recovery of Superquadrics, vol. 20. Springer, Dordrecht (2000). https://doi.org/10.1007/978-94-015-9456-1
3. Joshy, A.J., Hwang, J.T.: PySLSQP: a transparent python package for the SLSQP optimization algorithm modernized with utilities for visualization and post-processing. arXiv preprint arXiv:2408.13420 (2024)
4. Kraft, D.: A software package for sequential quadratic programming. Forschungsbericht- Deutsche Forschungs- und Versuchsanstalt fur. Luft- und Raumfahrt (1988)
5. Kumar, S., Hanson, R.: Robust methods for estimating pose and recovering shape from range data. Comput. Vis. Image Underst. **61**(3), 313–337 (1995). https://doi.org/10.1006/cviu.1995.1024

6. Makhal, A., Thomas, F., Gracia, A.P.: Grasping unknown objects in clutter by superquadric representation. In: Second IEEE International Conference on Robotic Computing (IRC), pp. 292–299. IEEE (2018)
7. Portal, R., Ferreira, A., Macho, A., Lima, P.U.: Using superquadrics for efficient collision detection in robotic manipulation. Robotica **28**(2), 253–265 (2010). https://doi.org/10.1017/S0263574709990519
8. Costa e Silva, E., Costa, M., Erlhagen, W., Bicho, E.: Superquadrics objects representation for robot manipulation. AIP Conf. Proc. **1738**, 300004 (2016)
9. Šircelj, J., et al.: Segmentation and recovery of superquadric models using convolutional neural networks. arXiv preprint arXiv:2001.10504 (2020)
10. Solina, F., Bajcsy, R.: Recovery of parametric models from range images: the case for superquadrics with global deformations. IEEE Trans. Pattern Anal. Mach. Intell. **12**(2), 131–147 (1990). https://doi.org/10.1109/34.45763
11. Torres, J.B., Carmona, E., Makhal, A., Heidari, O., Gracia, A.P.: Supertoroid fitting of objects with holes for robotic grasping and scene generation. arXiv preprint arXiv:2412.04174 (2024)

On the Dependency of Neural Differential Equations on the Training Numerical Method: Effects on Optimisation and Predictive Performance

C. Coelho(✉) and O. Niggemann

Institute for Artificial Intelligence, Helmut Schmidt University, Hamburg 22043, Germany
cecilia.coelho@hsu-hh.de

Abstract. Neural Ordinary Differential Equations (Neural ODEs) model data as the solution of a parametrised ordinary differential equation, evaluated numerically at observation times. In practice, both training and prediction require the choice of a numerical method and its hyperparameters. While classical numerical analysis guarantees that all consistent methods converge to the same continuous solution as the step size tends to zero, Neural ODEs are trained and evaluated at finite resolution, where this choice becomes an integral part of the model itself. In this work, we theoretically show that the loss function optimised during training is a surrogate loss determined by the chosen numerical method and its configuration. Under standard regularity assumptions on the dynamics and the loss, we derive explicit bounds on how the order and step size of the integrator perturb the loss surface, its gradient, and the location of the resulting minimisers. These bounds quantify the discrepancy between training and testing losses when different solvers are used, and reveal that method-induced errors behave as a systematic model misspecification term. We validate these theory by providing numerical experiments on a popular spiral ODE benchmark, where we observe large train-test discrepancies and qualitatively different optimisation paths across solvers, even when training errors are nearly identical. Furthermore, the theoretical analysis developed in this work applies to any neural network architecture whose output is defined via a numerical solution of a differential equation, including Neural ODEs, continuous normalizing flows, and related models. This work calls for the attention to, and explicit reporting of, the numerical solvers and their configurations in the Neural ODE literature and in applications that rely on these models.

Keywords: Neural Networks · Neural ode · Continuous-normalizing flows · Differential Equations · Numerical Methods · Numerical Analysis · Optimisation

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 255–271, 2027.
https://doi.org/10.1007/978-3-032-30530-5_16

1 Introduction

Neural Ordinary Differential Equations (Neural ODEs) are an increasingly popular neural network architecture that models ordinary differential equations (ODEs) to data. This allows to have continuous-time representations of datasets, previously impossible with classical neural networks [3]. Neural ODEs have been applied to various domains, including physics [13,21], chemistry [12,16], engineering [4,20], economics [11,14], and biology [17,18].

Furthermore, given the advantages and seamless modelling of systems that are classically modelled using differential equations, several approaches have branched out by improving on the basic ODEs to more complex differential equations, such as Neural Stochastic Differential Equations [10], Neural Delay Differential Equations [22], Neural Fractional Differential Equations [7], and Neural Chronos ODE [5]. Furthermore, Neural ODEs have been used for generative models, such as Continuous Normalizing Flows [3], and Latent ODEs [19].

In all of these approaches, the continuous dynamics are never accessed directly, both training and prediction rely on a numerical method tailored to the specific differential equation type, together with its associated hyperparameters such as step size or error tolerances. Classical numerical analysis guarantees that, for a fixed vector field and under suitable conditions, all consistent methods converge to the same continuous-time solution as the step size tends to zero [8]. In theory, this suggests that the choice of integrator should be largely irrelevant. However, in practice, Neural ODEs are trained and evaluated at finite resolution, with a fixed solver and a finite step size, so that the numerical method becomes an integral part of the model.

In the literature, several applications of Neural ODEs (and related architectures) have been proposed. However, to the best of our knowledge, none explicitly emphasises that the numerical method used during training should be the same as the one used for prediction. This omission can become problematic since domain experts, such as biologists or medical researchers, may not be aware of the subtle numerical dependence, which can lead to seemingly failed implementations and, in turn, distrust in Neural ODEs and related machine learning methods. More broadly, there is currently a gap between the machine learning viewpoint, where discretisation is treated as a minor implementation detail, and the actual practice, where the discretisation is part of the learned model.

In this paper, we build on the preliminary study of [6], which first highlighted empirically that Neural ODE parameters depend on the training numerical method. However, [6] is limited in several respects: the theoretical argument was informal and tied to a specific absolute-error loss; the impact of the solver was only discussed at the level of prediction error, without analysing gradients or minimisers; and the empirical evidence was restricted to a simple regression task evaluating predictive performance only.

To address the aforementioned gap this paper makes the following contributions:

- *Theoretical characterisation of solver-induced loss bias.* Under standard regularity assumptions, we show that the loss function optimised during training is the loss of a surrogate determined by the chosen numerical method and its configuration. We derive bounds of order $\mathcal{O}(h^p)$ that relate the order p and step size h of the integrator to the discrepancy between the *ideal* loss and the method-dependent *effective* loss;
- *Train-test discrepancy for different numerical methods.* We quantify how much the loss changes when the numerical method used at prediction time differs from the one used during training. The derived bounds show explicitly how the orders and step sizes of two different methods control the train-test discrepancy, and show that solver choice acts as a systematic model misspecification term when changed at test time;
- *Gradient-level analysis and parameter displacement.* We prove that the gradient of the *effective* loss obtained by differentiating through the numerical solver is a biased approximation of the *ideal* gradient, with a bias that scales in $\mathcal{O}(h^p)$. Under a local strong convexity assumption, these bounds show that the minimisers associated with different numerical methods can be separated in parameter space by a distance controlled by their respective global errors;
- *Empirical validation of the proposed statement.* Through numerical experiments we prove our statement by varying the training and testing solvers, as well as by performing a fixed-seed study that isolates the effect of the solver on the optimisation path.

As our analysis and experiments will show, the numerical method and its configuration are not mere simple implementation details but part of the model itself. This calls for greater attention to, and explicit reporting of, the chosen solvers and their hyperparameters in the Neural ODE literature and in applications that rely on these models and related neural differential equation architectures.

Note that, although we herein focus on Neural ODEs, the theoretical analysis applies to any architecture whose outputs are obtained by numerically solving an underlying dynamical system, including continuous normalizing flows and related neural differential equation models.

The rest of this paper is organised as follows. Section 2 briefly introduces the necessary concepts on Neural ODEs, and numerical methods. Section 3 develops our theoretical analysis of how the choice of numerical method affects the effective loss, gradients, and learnt parameters of Neural ODE models. Section 4 presents numerical experiments on a controlled benchmark that illustrate and validate the theory. Finally, the paper ends with the conclusions in Sect. 5.

2 Background

2.1 Neural Ordinary Differential Equations

A Neural ODE is a neural network architecture that fits a network $\boldsymbol{f_\theta}$, with learnable parameters $\boldsymbol{\theta} \in \Theta$, to the right-hand side of an ODE, or system of

ODEs. The end goal is that solving this ODE as an initial value problem (IVP) fits the data:

$$\frac{d\boldsymbol{y}}{dt}(t) = \boldsymbol{f}_{\boldsymbol{\theta}}(t, \boldsymbol{y}(t)) \qquad \boldsymbol{y}(t_0) = \boldsymbol{y}_0, \tag{1}$$

where $t \in [t_0, t_M]$ and $\boldsymbol{y}(t) \in \mathbb{R}^d$ [3].

To obtain the predictions $\{\hat{\boldsymbol{y}}_m\}_{m=1}^M$, (1) needs to be solved numerically using a classical ODE solver:

$$\{\hat{\boldsymbol{y}}_m\}_{m=1}^M = \text{ODESolve}(\boldsymbol{f}_{\boldsymbol{\theta}}, \boldsymbol{y}_0, (t_0, t_M)).$$

These predictions are then used in a loss functional to optimise $\boldsymbol{\theta}$ by computing the error of the fit to the data [3].

2.2 Numerical Methods

Consider the classical ODE IVP:

$$\frac{d\boldsymbol{y}}{dt}(t) = \boldsymbol{f}(t, \boldsymbol{y}(t)) \qquad \boldsymbol{y}(t_0) = \boldsymbol{y}_0,$$

on $[t_0, t_M]$ and a discretisation grid $t_0 < t_1 < \cdots < t_M$, $t_{m+1} - t_m = h$, with constant step size $h > 0$, for simplicity. A (one-step) numerical method $\mathcal{A}$ computes an approximation $\boldsymbol{y}_m^{\mathcal{A}} \approx \boldsymbol{y}(t_m)$ by applying an update rule starting from $\boldsymbol{y}_0$. Different update rules lead to different solvers, with the most well known being: *Explicit Euler* (order $p = 1$) $\boldsymbol{y}_{m+1} = \boldsymbol{y}_m + h\boldsymbol{f}(t_m, \boldsymbol{y}_m)$; *explicit midpoint / Heun* (order $p = 2$), which uses an additional intermediate evaluation of $\boldsymbol{f}$ to improve accuracy; *runge-kutta* (*rk4*, order $p = 4$), which combines four stages per step; *dormand-prince* (*dopri5*), an adaptive runge-kutta method that adjusts h based on an embedded error estimate [1].

In the context of Neural ODEs, these numerical methods are used both in the forward pass to compute the predictions, and in the backward pass to compute gradients via backpropagation or the adjoint method [3].

3 Dependence on the Numerical Method

We now formalise how the numerical method enters the effective loss optimised by a Neural ODE and how changing the method between training and prediction perturbs this loss, thus affecting the resulting model.

For simplicity, consider a Neural ODE with a single trajectory and let $\{(t_m, \boldsymbol{y}_m)\}_{m=0}^M$ be observed data points, with $\boldsymbol{y}_0$ the initial condition at time t_0. For each parameter vector $\boldsymbol{\theta}$, the continuous dynamics are given by (1) with flow $\varphi_{\boldsymbol{\theta}}$.

We introduce the following assumptions:

A1) *Regularity of the vector field.*
each $\boldsymbol{\theta} \in \Theta$, the function $\boldsymbol{f_\theta}$ is globally Lipschitz in $\boldsymbol{y}$ and continuous in t on $[t_0, t_M] \times \mathbb{R}^d$, with Lipschitz constant $L_f(\boldsymbol{\theta})$. Then the IVP (1) has a unique solution on $[t_0, t_M]$, and the flow $\varphi_{\boldsymbol{\theta}}$ is well defined [8].
A2) *Numerical method, order, and global error.*
$\mathcal{A}$ denote a fixed-step numerical method of order $p \in \mathbb{N}$ with step size $h > 0$, used to integrate (1) from t_0 to t_m. We denote by $\Phi^{\mathcal{A}}_{\boldsymbol{\theta}}(t_m, t_0, \boldsymbol{y}_0)$ the resulting numerical approximation to $\varphi_{\boldsymbol{\theta}}(t_m, t_0, \boldsymbol{y}_0)$. Following classical numerical analysis for stable one-step methods, we assume that there exists a constant $C(\boldsymbol{\theta})$, independent of h and m, such that the global error satisfies:

$$\left\|\Phi^{\mathcal{A}}_{\boldsymbol{\theta}}(t_m, t_0, \boldsymbol{y}_0) - \varphi_{\boldsymbol{\theta}}(t_m, t_0, \boldsymbol{y}_0)\right\| \leq C(\boldsymbol{\theta})h^p, \qquad \forall m = 1, \ldots, M, \tag{2}$$

where the order p appears explicitly as the exponent governing how fast the numerical solution converges to the true flow as $h \to 0$ [9].
A3) *Lipschitz loss on bounded trajectories.*
the discrepancy between predictions and observations to be measured via a pointwise loss $\rho : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$. Then, assuming there exists $R > 0$ such that all true and predicted states satisfy $\|\boldsymbol{y}_m\| \leq R$ and $\|\hat{\boldsymbol{y}}_m(\boldsymbol{\theta}; \mathcal{A})\| \leq R$ for all relevant $\boldsymbol{\theta}$ and m. Then there exists $L_\rho(R) > 0$ such that, for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{y} \in \mathbb{R}^d$ with $\|\boldsymbol{u}\|, \|\boldsymbol{v}\|, \|\boldsymbol{y}\| \leq R$,

$$|\rho(\boldsymbol{u}, \boldsymbol{y}) - \rho(\boldsymbol{v}, \boldsymbol{y})| \leq L_\rho(R)\|\boldsymbol{u} - \boldsymbol{v}\|.$$

This condition is mild and standard in error-propagation arguments, *i.e.* if $\rho(\cdot, \boldsymbol{y})$ is continuously differentiable then one can take $L_\rho(R) = sup_{\|\boldsymbol{u}\|, \|\boldsymbol{y}\|} \leq R\|\lambda_{\boldsymbol{u}}\rho(\boldsymbol{u}, \boldsymbol{y})\|$.
Note that, the boundeness requirement on trajectories can be ensured in practise by restricting the data range using standard regularisation or applying mild growth controls on $\boldsymbol{f_\theta}$ over finite time intervals.

Given $\boldsymbol{\theta}$, consider the *ideal* loss function:

$$\mathcal{L}^*(\boldsymbol{\theta}) = \frac{1}{M}\sum_{m=1}^{M} \rho\big(\hat{\boldsymbol{y}}^*_m(\boldsymbol{\theta}), \boldsymbol{y}_m\big), \qquad \hat{\boldsymbol{y}}^*_m(\boldsymbol{\theta}) = \varphi_{\boldsymbol{\theta}}(t_m, t_0, \boldsymbol{y}_0).$$

In practice, the numerical method $\mathcal{A}$ is used to compute predictions, $\hat{\boldsymbol{y}}_m(\boldsymbol{\theta}; \mathcal{A}) = \Phi^{\mathcal{A}}_{\boldsymbol{\theta}}(t_m, t_0, \boldsymbol{y}_0)$, so the *effective* training loss is instead given by:

$$\mathcal{L}^{\mathcal{A}}(\boldsymbol{\theta}) = \frac{1}{M}\sum_{m=1}^{M} \rho\big(\hat{\boldsymbol{y}}_m(\boldsymbol{\theta}; \mathcal{A}), \boldsymbol{y}_m\big). \tag{3}$$

Thus, given the dependence of $\hat{\boldsymbol{y}}$ on $\mathcal{A}$, the numerical method defines the actual objective surface seen by the optimiser.

Proposition 1 (Loss perturbation induced by a numerical method). *Suppose Assumptions A1, A2, and A3 hold. Then for every* $\boldsymbol{\theta} \in \Theta$,

$$\left|\mathcal{L}^{\mathcal{A}}(\boldsymbol{\theta}) - \mathcal{L}^{*}(\boldsymbol{\theta})\right| \leq L_{\rho}(R)C(\boldsymbol{\theta})h^{p}. \tag{4}$$

Note that, by A2 the constant $C(\boldsymbol{\theta})$ *is uniform in the time index* m *(and independent of* h*), so the bound* $\left\|\hat{\boldsymbol{y}}_m(\boldsymbol{\theta};\mathcal{A}) - \hat{\boldsymbol{y}}^{*}_m(\boldsymbol{\theta})\right\| \leq C(\boldsymbol{\theta})h^{p}$ *holds for every summand. Thus, averaging over* m *leaves the right-hand side unchanged.*

Proof. Consider a fix set of network parameters $\boldsymbol{\theta}_1 \in \Theta$. Then, by Assumption A2, for all $m \in 1, \ldots, M$,

$$\|\hat{\boldsymbol{y}}_m(\boldsymbol{\theta};\mathcal{A}) - \hat{\boldsymbol{y}}^{*}_m(\boldsymbol{\theta})\| = \left\|\Phi^{\mathcal{A}}_{\boldsymbol{\theta}}(t_m, t_0, \boldsymbol{y}_0) - \varphi_{\boldsymbol{\theta}}(t_m, t_0, \boldsymbol{y}_0)\right\| \leq C(\boldsymbol{\theta})h^{p}.$$

Using Assumption A3, and considering $\boldsymbol{u} = \hat{\boldsymbol{y}}_m(\boldsymbol{\theta};\mathcal{A})$, $\boldsymbol{v} = \hat{\boldsymbol{y}}^{*}_m(\boldsymbol{\theta})$, and $\boldsymbol{y} = \boldsymbol{y}_m$, comes:

$$\left|\rho\big(\hat{\boldsymbol{y}}_m(\boldsymbol{\theta};\mathcal{A}), \boldsymbol{y}_m\big) - \rho\big(\hat{\boldsymbol{y}}^{*}_m(\boldsymbol{\theta}), \boldsymbol{y}_m\big)\right| \leq L_{\rho}(R)\left\|\hat{\boldsymbol{y}}_m(\boldsymbol{\theta};\mathcal{A}) - \hat{\boldsymbol{y}}^{*}_m(\boldsymbol{\theta})\right\| \leq L_{\rho}(R)C(\boldsymbol{\theta})h^{p}.$$

Averaging over all m data points results:

$$\begin{aligned}
\left|\mathcal{L}^{\mathcal{A}}(\boldsymbol{\theta}) - \mathcal{L}^{*}(\boldsymbol{\theta})\right| &= \left|\frac{1}{M}\sum_{m=1}^{M}\left(\rho\big(\hat{\boldsymbol{y}}_m(\boldsymbol{\theta};\mathcal{A}), \boldsymbol{y}_m\big) - \rho\big(\hat{\boldsymbol{y}}^{*}_m(\boldsymbol{\theta}), \boldsymbol{y}_m\big)\right)\right| \\
&\leq \frac{1}{M}\sum_{m=1}^{M} L_{\rho}(R)C(\boldsymbol{\theta})h^{p} \\
&= L_{\rho}(R)C(\boldsymbol{\theta})h^{p},
\end{aligned}$$

which is (4).

Proposition 1 shows that the numerical method perturbs the loss by an amount of order $\mathcal{O}(h^p)$. For a fixed h, this perturbation is systematic, meaning that even with infinite data and perfect optimisation, the parameters $\boldsymbol{\theta}$ are chosen to minimise $\mathcal{L}^{\mathcal{A}}$, not $\mathcal{L}^{*}$. The factor h^p in (4) shows how quickly the solver-induced bias decays as the step size $h \to 0$. For a first-order method (p=1, such as the explicit Euler), halving the step size halves the error term, since it scales in h. For a second-order method (p=2), the same refinement reduces the error by a factor of four, since it scales in h^2. For a fourth-order method (p=4, such as the Runge-Kutta), halving h reduces the error by a factor of sixteen, as it scales in h^4 [1]. Thus, higher-order methods make the discrepancy between the *ideal* $\mathcal{L}^{*}$ and *effective* $\mathcal{L}^{\mathcal{A}}$ losses vanish much faster as the time discretisation is refined.

3.1 Training Versus Prediction with Different Methods

We now compare the usage of two different numerical methods, one during training and another during prediction.

Let $\mathcal{A}_1$ and $\mathcal{A}_2$ denote two numerical methods of orders p_1 and p_2 with step sizes h_1 and h_2, respectively. Then, the induced numerical flows $\Phi_{\boldsymbol{\theta}}^{\mathcal{A}_1}$ and $\Phi_{\boldsymbol{\theta}}^{\mathcal{A}_2}$, and corresponding losses $\mathcal{L}^{\mathcal{A}_1}$ and $\mathcal{L}^{\mathcal{A}_2}$ can be defined as in (3).

Let us assume that each method satisfies its own global error bound given by [9]:

$$\left\|\Phi_{\boldsymbol{\theta}}^{\mathcal{A}_1}(t_m, t_0, \boldsymbol{y}_0) - \varphi_{\boldsymbol{\theta}}(t_m, t_0, \boldsymbol{y}_0)\right\| \leq C_1(\boldsymbol{\theta}) h_1^{p_1}, \tag{5}$$

$$\left\|\Phi_{\boldsymbol{\theta}}^{\mathcal{A}_2}(t_m, t_0, \boldsymbol{y}_0) - \varphi_{\boldsymbol{\theta}}(t_m, t_0, \boldsymbol{y}_0)\right\| \leq C_2(\boldsymbol{\theta}) h_2^{p_2}. \tag{6}$$

Proposition 2 (Discrepancy between training and testing losses). *Under Assumptions A1, A2 and A3, and the error bounds* (5) *and* (6)*, for every* $\boldsymbol{\theta} \in \Theta$ *comes:*

$$\left|\mathcal{L}^{\mathcal{A}_2}(\boldsymbol{\theta}) - \mathcal{L}^{\mathcal{A}_1}(\boldsymbol{\theta})\right| \leq L_\rho(R)\left(C_1(\boldsymbol{\theta}) h_1^{p_1} + C_2(\boldsymbol{\theta}) h_2^{p_2}\right), \tag{7}$$

and when the same method is used for training and testing ($\mathcal{A}_1 = \mathcal{A}_2$)*, comes* $\mathcal{L}^{\mathcal{A}_1}(\boldsymbol{\theta}) = \mathcal{L}^{\mathcal{A}_2}(\boldsymbol{\theta})$*, for all* $\boldsymbol{\theta} \in \Theta$*. Any remaining discrepancy is then solely the bias of* $\mathcal{L}^{\mathcal{A}}$ *with respect to* $\mathcal{L}^*$*, which is of order* $\mathcal{O}(h^p)$ *by Proposition 1.*

Proof. Using the triangle inequality [9], and adding and subtracting the losses,

$$\left|\mathcal{L}^{\mathcal{A}_2}(\boldsymbol{\theta}) - \mathcal{L}^{\mathcal{A}_1}(\boldsymbol{\theta})\right| \leq \left|\mathcal{L}^{\mathcal{A}_2}(\boldsymbol{\theta}) - \mathcal{L}^*(\boldsymbol{\theta})\right| + \left|\mathcal{L}^{\mathcal{A}_1}(\boldsymbol{\theta}) - \mathcal{L}^*(\boldsymbol{\theta})\right|. \tag{8}$$

Applying Proposition 1 to $\mathcal{A}_1$ and $\mathcal{A}_2$ separately, with constants $C_1(\boldsymbol{\theta})$ and $C_2(\boldsymbol{\theta})$, comes:

$$\left|\mathcal{L}^{\mathcal{A}_2}(\boldsymbol{\theta}) - \mathcal{L}^*(\boldsymbol{\theta})\right| \leq L_\rho(R) C_2(\boldsymbol{\theta}) h_2^{p_2}, \qquad \left|\mathcal{L}^{\mathcal{A}_1}(\boldsymbol{\theta}) - \mathcal{L}^*(\boldsymbol{\theta})\right| \leq L_\rho(R) C_1(\boldsymbol{\theta}) h_1^{p_1},$$

and by substitution of these inequalities in (8) we reach to (7).

Proposition 2 makes it explicit that changing the numerical method between training and prediction perturbs the loss landscape by an amount controlled by the global errors and by the orders p_1 and p_2 of both methods. Furthermore, changing the associated step sizes h_1 and h_2 also affects the magnitude of this perturbation through the factors $h_1^{p_1}$ and $h_2^{p_2}$. Also, changing to lower order methods (small p) or larger step sizes h lead to larger perturbations.

Moreover, for any fixed $\boldsymbol{\theta}$, the difference in predictions produced by the two numerical methods can be bounded directly using (5) and (6):

$$\left\|\hat{\boldsymbol{y}}_m(\boldsymbol{\theta}; \mathcal{A}_2) - \hat{\boldsymbol{y}}_m(\boldsymbol{\theta}; \mathcal{A}_1)\right\| \leq C_1(\boldsymbol{\theta}) h_1^{p_1} + C_2(\boldsymbol{\theta}) h_2^{p_2}, \qquad \forall m = 1, \ldots, M.$$

Thus, even with the same parameters $\boldsymbol{\theta}$, changing the solver to one of different order changes the outputs by a computable amount depending on $h_1^{p_1}$ and $h_2^{p_2}$.

3.2 Gradient-Level Effects of the Numerical Method

During training, the Neural ODE parameter updates are driven by the gradient of the *effective* loss, $\nabla_{\boldsymbol{\theta}} \mathcal{L}^{\mathcal{A}}(\boldsymbol{\theta})$, which is obtained by differentiating through the numerical solver (using backpropagation or the adjoint method [3]). We now show that, under Assumptions A1-A3 and some additinal assumptions, $\nabla_{\boldsymbol{\theta}} \mathcal{L}^{\mathcal{A}}(\boldsymbol{\theta})$ is a perturbation of the *ideal* loss gradient $\nabla_{\boldsymbol{\theta}} \mathcal{L}^{*}(\boldsymbol{\theta})$ of order $\mathcal{O}(h^p)$, with an explicit dependence on the solver order p and step size h.

We impose the following additional assumptions:

A4) *Smoothness in parameters and states.*
each $t \in [t_0, t_M]$, the vector field $\boldsymbol{f}_{\boldsymbol{\theta}}(t, \boldsymbol{y})$ is continuously differentiable in both $\boldsymbol{\theta}$ and $\boldsymbol{y}$, with bounded derivatives on compact sets. Moreover, the pointwise loss $\rho(\boldsymbol{u}, \boldsymbol{y})$ is continuously differentiable in $\boldsymbol{u}$ with bounded gradient.
Under this assumption, the sensitivity of the continuous flow with respect to the parameters,

$$S_{\boldsymbol{\theta}}(t) = \frac{\partial \varphi_{\boldsymbol{\theta}}(t, t_0, \boldsymbol{y}_0)}{\partial \boldsymbol{\theta}},$$

exists and solves the linear ODE, $\dot{S}_{\boldsymbol{\theta}}(t) = \dfrac{\partial \boldsymbol{f}_{\boldsymbol{\theta}}}{\partial \boldsymbol{y}}(t, \boldsymbol{y}(t)) S_{\boldsymbol{\theta}}(t) + \dfrac{\partial \boldsymbol{f}_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}}(t, \boldsymbol{y}(t))$,
with initial condition $S_{\boldsymbol{\theta}}(t_0) = 0$.
A5) *Global error for sensitivities.*
to the state variables, the numerical method $\mathcal{A}$ induces a discrete sensitivity:

$$S^{\mathcal{A}}_{\boldsymbol{\theta}}(t_m) = \frac{\partial \Phi^{\mathcal{A}}_{\boldsymbol{\theta}}(t_m, t_0, \boldsymbol{y}_0)}{\partial \boldsymbol{\theta}}, \qquad \forall m = 1, \ldots, M$$

which is the quantity used by automatic differentiation. Then we assume that there exists a constant $C_S(\boldsymbol{\theta})$, independent of h and m, such that [9]:

$$\left\| S^{\mathcal{A}}_{\boldsymbol{\theta}}(t_m) - S_{\boldsymbol{\theta}}(t_m) \right\| \leq C_S(\boldsymbol{\theta}) h^p.$$

Proposition 3 (Gradient perturbation induced by the numerical method). *Suppose Assumptions A1-A5 hold, then for every $\boldsymbol{\theta} \in \Theta$,*

$$\left\| \nabla_{\boldsymbol{\theta}} \mathcal{L}^{\mathcal{A}}(\boldsymbol{\theta}) - \nabla_{\boldsymbol{\theta}} \mathcal{L}^{*}(\boldsymbol{\theta}) \right\| \leq K(\boldsymbol{\theta}) h^p, \tag{9}$$

for some constant $K(\boldsymbol{\theta})$ depending on the bounds for $\boldsymbol{f}_{\boldsymbol{\theta}}$, its derivatives, and the loss ρ, but independent of h and m. Thus, the training gradient obtained by differentiating through the numerical solver is a biased approximation of the ideal *gradient, with bias of order $\mathcal{O}(h^p)$.*

Proof. By the chain rule, the *ideal* gradient can be written as:

$$\nabla_{\boldsymbol{\theta}} \mathcal{L}^{*}(\boldsymbol{\theta}) = \frac{1}{M} \sum_{m=1}^{M} \Big(D_{\boldsymbol{u}} \rho\big(\hat{\boldsymbol{y}}^{*}_m(\boldsymbol{\theta}), \boldsymbol{y}_m\big) \Big) S_{\boldsymbol{\theta}}(t_m), \tag{10}$$

where $D_{\boldsymbol{u}} \rho$ is the gradient of ρ with respect to its first component.

Similarly, the *effective* gradient can be expressed as:

$$\nabla_{\boldsymbol{\theta}} \mathcal{L}^{\mathcal{A}}(\boldsymbol{\theta}) = \frac{1}{M} \sum_{m=1}^{M} \Big(D_{\boldsymbol{u}} \rho\big(\hat{\boldsymbol{y}}_m(\boldsymbol{\theta}; \mathcal{A}), \boldsymbol{y}_m\big) \Big) S_{\boldsymbol{\theta}}^{\mathcal{A}}(t_m). \tag{11}$$

Subtracting (10) and (11) comes:

$$\nabla_{\boldsymbol{\theta}} \mathcal{L}^{\mathcal{A}}(\boldsymbol{\theta}) - \nabla_{\boldsymbol{\theta}} \mathcal{L}^{*}(\boldsymbol{\theta}) = \frac{1}{M} \sum_{m=1}^{M} \Big[\Delta_m^{(1)}(\boldsymbol{\theta}) + \Delta_m^{(2)}(\boldsymbol{\theta}) \Big],$$

with $\Delta_m^{(1)}(\boldsymbol{\theta}) = \Big(D_{\boldsymbol{u}} \rho\big(\hat{\boldsymbol{y}}_m(\boldsymbol{\theta}; \mathcal{A}), \boldsymbol{y}_m\big) - D_{\boldsymbol{u}} \rho\big(\hat{\boldsymbol{y}}_m^{*}(\boldsymbol{\theta}), \boldsymbol{y}_m\big) \Big) S_{\boldsymbol{\theta}}(t_m)$, and $\Delta_m^{(2)}(\boldsymbol{\theta}) = D_{\boldsymbol{u}} \rho\big(\hat{\boldsymbol{y}}_m(\boldsymbol{\theta}; \mathcal{A}), \boldsymbol{y}_m\big) \big(S_{\boldsymbol{\theta}}^{\mathcal{A}}(t_m) - S_{\boldsymbol{\theta}}(t_m) \big)$.

From Assumption A4 we have that $D_{\boldsymbol{u}} \rho$ is Lipschitz in $\boldsymbol{u}$, and the sensitivity $S_{\boldsymbol{\theta}}(t_m)$ is bounded. Thus, there exists a constant $C_{S_1}(\boldsymbol{\theta}) > 0$ such that:

$$\big\| \Delta_m^{(1)}(\boldsymbol{\theta}) \big\| \leq C_{S_1}(\boldsymbol{\theta}) \big\| \hat{\boldsymbol{y}}_m(\boldsymbol{\theta}; \mathcal{A}) - \hat{\boldsymbol{y}}_m^{*}(\boldsymbol{\theta}) \big\|.$$

And by combing with the global error bound (2) comes:

$$\big\| \Delta_m^{(1)}(\boldsymbol{\theta}) \big\| \leq C_{S_1}(\boldsymbol{\theta}) C(\boldsymbol{\theta}) h^p.$$

Similarly, from Assumption A5 comes:

$$\| \Delta_m^{(2)}(\boldsymbol{\theta}) \| \leq C_{S_2}(\boldsymbol{\theta}) \| S_{\boldsymbol{\theta}}^{\mathcal{A}}(t_m) - S_{\boldsymbol{\theta}}(t_m) \| \leq C_{S_2}(\boldsymbol{\theta}) C_S(\boldsymbol{\theta}) h^p,$$

and by combining these bounds and averaging over all m we reach to (9), with $K(\boldsymbol{\theta}) = C_{S_1}(\boldsymbol{\theta}) C(\boldsymbol{\theta}) + C_{S_2}(\boldsymbol{\theta}) C_S(\boldsymbol{\theta})$.

Proposition 3 formalises that the choice of numerical method, its step size h, and its order p does not only perturb the value of the loss, but also the gradient directions and magnitudes during optimisation. Consequently, training with different numerical integrators is equivalent to performing gradient descent on different, method-dependent approximations of the optimisation objective, where the approximation error scales with h^p.

3.3 Consequences for the Learnt Parameters and Results

Let $\boldsymbol{\theta}_1^{\star} \in \arg\min_{\boldsymbol{\theta} \in \Theta} \mathcal{L}^{\mathcal{A}_1}(\boldsymbol{\theta})$ be a local minimiser of the training loss associated with method $\mathcal{A}_1$. Then, from Proposition 2 comes that the testing loss evaluated with a different method satisfies:

$$\mathcal{L}^{\mathcal{A}_2}(\boldsymbol{\theta}_1^{\star}) \leq \mathcal{L}^{\mathcal{A}_1}(\boldsymbol{\theta}_1^{\star}) + L_{\rho}(R) \big(C_1(\boldsymbol{\theta}_1^{\star}) h_1^{p_1} + C_2(\boldsymbol{\theta}_1^{\star}) h_2^{p_2} \big).$$

Thus, even if $\boldsymbol{\theta}_1^{\star}$ is nearly optimal for $\mathcal{L}^{\mathcal{A}_1}$, it may be sub-optimal for $\mathcal{L}^{\mathcal{A}_2}$, with a difference controlled by the orders p_1 and p_2, and by the respective step sizes.

To clearly show the impact on the optimisation process, we now relate the numerical method to the location of the minimisers in the parameter

space. Assume that $\mathcal{L}^*$ is locally μ-strongly convex around a minimiser $\boldsymbol{\theta}^\star \in \arg\min_{\boldsymbol{\theta}} \mathcal{L}^*(\boldsymbol{\theta})$, in the sense of [15, Definition 2.1.2], and that the gradient perturbation bounds hold as in Proposition 3. Let $\boldsymbol{\theta}_1^\star \in \arg\min_{\boldsymbol{\theta}} \mathcal{L}^{\mathcal{A}_1}(\boldsymbol{\theta})$ and $\boldsymbol{\theta}_2^\star \in \arg\min_{\boldsymbol{\theta}} \mathcal{L}^{\mathcal{A}_2}(\boldsymbol{\theta})$. Then, by Theorem 2.1.9 in [15], strong convexity implies that for any $\boldsymbol{\theta}$ in a neighbourhood of $\boldsymbol{\theta}^\star$ we have:

$$\left\langle \nabla_{\boldsymbol{\theta}} \mathcal{L}^*(\boldsymbol{\theta}) - \nabla_{\boldsymbol{\theta}} \mathcal{L}^*(\boldsymbol{\theta}^\star), \boldsymbol{\theta} - \boldsymbol{\theta}^\star \right\rangle \;\geq\; \mu \|\boldsymbol{\theta} - \boldsymbol{\theta}^\star\|^2,$$

and since $\nabla_{\boldsymbol{\theta}} \mathcal{L}^*(\boldsymbol{\theta}^\star) = \mathbf{0}$, then it reduces to:

$$\left\langle \nabla_{\boldsymbol{\theta}} \mathcal{L}^*(\boldsymbol{\theta}), \boldsymbol{\theta} - \boldsymbol{\theta}^\star \right\rangle \;\geq\; \mu \|\boldsymbol{\theta} - \boldsymbol{\theta}^\star\|^2.$$

Applying the Cauchy-Schwarz inequality and considering $\boldsymbol{\theta} \neq \boldsymbol{\theta}^\star$:

$$\mu \|\boldsymbol{\theta} - \boldsymbol{\theta}^\star\| \;\leq\; \left\| \nabla_{\boldsymbol{\theta}} \mathcal{L}^*(\boldsymbol{\theta}) \right\|.$$

By taking $\boldsymbol{\theta} = \boldsymbol{\theta}_1^\star$ and using that $\nabla_{\boldsymbol{\theta}} \mathcal{L}^{\mathcal{A}_1}(\boldsymbol{\theta}_1^\star) = \mathbf{0}$, we get:

$$\mu \|\boldsymbol{\theta}_1^\star - \boldsymbol{\theta}^\star\| \leq \left\| \nabla_{\boldsymbol{\theta}} \mathcal{L}^*(\boldsymbol{\theta}_1^\star) \right\| = \left\| \nabla_{\boldsymbol{\theta}} \mathcal{L}^*(\boldsymbol{\theta}_1^\star) - \nabla_{\boldsymbol{\theta}} \mathcal{L}^{\mathcal{A}_1}(\boldsymbol{\theta}_1^\star) \right\| \leq K_1(\boldsymbol{\theta}_1^\star) h_1^{p_1},$$

and therefore

$$\|\boldsymbol{\theta}_1^\star - \boldsymbol{\theta}^\star\| \leq \frac{K_1}{\mu} h_1^{p_1}.$$

Following an analogous argument for $\mathcal{A}_2$ comes:

$$\|\boldsymbol{\theta}_2^\star - \boldsymbol{\theta}^\star\| \leq \frac{K_2}{\mu} h_2^{p_2}.$$

Finally, using the triangle inequality, we reach:

$$\|\boldsymbol{\theta}_1^\star - \boldsymbol{\theta}_2^\star\| \leq \|\boldsymbol{\theta}_1^\star - \boldsymbol{\theta}^\star\| + \|\boldsymbol{\theta}^\star - \boldsymbol{\theta}_2^\star\| \lesssim \frac{K_1}{\mu} h_1^{p_1} + \frac{K_2}{\mu} h_2^{p_2}.$$

Thus, we conclude that lower order methods (smaller p) or larger step sizes can lead to minimisers that are significantly displaced from the ideal minimiser $\boldsymbol{\theta}^\star$ and from each other.

4 Numerical Experiments

To support the theoretical analysis in Sect. 3, we compare different solvers during training and testing on a popular benchmark ODE by using the Mean Squared Error (MSE) as the loss and evaluation metric, and explicitly studying the effect of solver order on both the optimisation process and the final predictions.

We use a two-dimensional linear spiral ODE with an analytic solution, which allows us to separate discretisation error from optimisation effects. The ground-truth dynamics are given by:

$$\dot{\boldsymbol{y}}(t) = A\boldsymbol{y}(t), \text{ with } A = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}, \boldsymbol{y}(t) \in \mathbb{R}^2,$$

where $\alpha < 0$ controls radial decay and $\beta > 0$ controls rotational speed. In all experiments we fix $\alpha = -0.1$, $\beta = 1.0$, and the initial condition $y_0 = (2, 0)^\top$. The exact solution is:

$$\boldsymbol{y}(t) = \exp(tA)\boldsymbol{y}_0,$$

which we evaluate analytically. We generate a training/reconstruction dataset by sampling $N_{\text{train}} = 50$ time points on a uniform grid $t_i \in [0, T_{\text{train}}]$, $T_{\text{train}} = 5.0$, to get pairs $(t_i, \boldsymbol{y}(t_i))$. A test/extrapolation dataset is also generated by sampling $N_{\text{test}} = 100$ time points on a uniform grid $t_j \in [0, T_{\text{test}}]$, $T_{\text{test}} = 10.0$.

We use the same Neural ODE configuration across experiments, where $\boldsymbol{f_\theta}$ is a fully-connected neural network with one hidden layer with 32 neurons and a hyperbolic tangent activation function. We considered four ODE solvers: *euler* (explicit Euler, order $p = 1$); *midpoint* (explicit midpoint/Heun, order $p = 2$); *rk4* (Runge-Kutta, order $p = 4$); *dopri5* (adaptive Dormand–Prince method). For the fixed-step solvers (*euler*, *midpoint*, *rk4*) we use the same uniform time grid on $[0, T_{\text{train}}]$, so that they share a common step size h. The adaptive solver *dopri5* is run with its default tolerances (see [2]).

All models were trained with the Adam optimiser, using a learning rate 10^{-2} and 200 epochs, with batch size equal to the full training set. Each configuration is repeated with 3 random seeds and we report mean and standard deviation across seeds. All implementations were done using the *Torchdiffeq* library [2].

Note that, all Neural ODE models were trained by fixing seeds as to isolate the effect of changing the numerical method.

4.1 Effect of Different Solvers for Training and Testing

To study the impact of using different solvers during training and prediction, we consider a full grid of training and testing solvers. Let $\mathcal{S} = \{$*euler*, *midpoint*, *rk4*, *dopri5*$\}$. For each training solver $\mathcal{A}_{\text{tr}} \in \mathcal{S}$ and each random seed, we train a Neural ODE model by minimising $\mathcal{L}^{\mathcal{A}_{\text{tr}}}(\boldsymbol{\theta})$ on the training dataset, $[0, T_{\text{train}}]$ and we evaluate the trained model on the training grid using $\mathcal{A}_{\text{tr}}$ and compute the MSE against $\boldsymbol{y}(t)$ on $[0, T_{\text{train}}]$. Then, we evaluate the trained model on the training grid using all $\mathcal{A}_{\text{te}} \in \mathcal{S}$ and compute the MSE against $\boldsymbol{y}(t)$ on $[0, T_{\text{train}}]$. Finally, we do an extrapolation experiment by evaluating the same model on the extended test grid using all $\mathcal{A}_{\text{te}}$ and compute the MSE against $\boldsymbol{y}(t)$ on $[0, T_{\text{test}}]$.

Table 1, Table 2, and Table 3 present the reconstruction and extrapolation results, respectively, with the average MSE and its respective standard deviation (std), over the three seeds, for each pair $(\mathcal{A}_{\text{tr}}, \mathcal{A}_{\text{te}})$. Figure 1 and Fig. 2 show the predictions of the Neural ODE models trained with $\mathcal{A}_{\text{tr}}$ and using $\mathcal{A}_{\text{te}}$ for the reconstruction and extrapolation experiments, respectively.

Table 1. Training error (mean MSE ± std) for different training (row) and testing (column) solvers.

Training	Testing			
	dopri5	euler	midpoint	rk4
dopri5	1.62e-03 ± 3.56e-04	1.62e-03 ± 3.56e-04	1.62e-03 ± 3.56e-04	1.62e-03 ± 3.56e-04
euler	1.38e-03 ± 2.83e-04	1.38e-03 ± 2.83e-04	1.38e-03 ± 2.83e-04	1.38e-03 ± 2.83e-04
midpoint	1.63e-03 ± 3.61e-04	1.63e-03 ± 3.61e-04	1.63e-03 ± 3.61e-04	1.63e-03 ± 3.61e-04
rk4	1.62e-03 ± 3.56e-04	1.62e-03 ± 3.56e-04	1.62e-03 ± 3.56e-04	1.62e-03 ± 3.56e-04

Table 2. Reconstruction error (mean MSE ± std) for different training (row) and testing (column) solvers.

Training	Testing			
	dopri5	euler	midpoint	rk4
dopri5	1.62e-03 ± 3.56e-04	2.71e-02 ± 3.23e-03	1.65e-03 ± 3.54e-04	1.62e-03 ± 3.56e-04
euler	1.94e-02 ± 2.23e-03	1.38e-03 ± 2.83e-04	2.01e-02 ± 2.16e-03	1.94e-02 ± 2.23e-03
midpoint	1.63e-03 ± 3.63e-04	2.78e-02 ± 3.14e-03	1.63e-03 ± 3.61e-04	1.63e-03 ± 3.63e-04
rk4	1.62e-03 ± 3.56e-04	2.71e-02 ± 3.23e-03	1.65e-03 ± 3.54e-04	1.62e-03 ± 3.56e-04

From Table 1, we see that a Neural ODE trained with *euler* achieves the lowest training error, while the higher-order solvers *dopri*, *rk4*, and *midpoint* get slightly larger values, showing that all solvers can fit the data on the training grid to essentially the same level, with small variability across seeds.

In Table 2, the diagonal entries coincide with the training errors, as expected when training and testing solvers agree and the training and testing datasets are the same. Off-diagonal entries, however, can be an order of magnitude larger, especially when testing with *euler* a model trained with a higher-order solver, or vice versa. In conjunction with Fig. 1, this directly illustrates the train-test discrepancy showed in our theoretical analysis, Sect. 3, in which different solvers induce different effective loss landscapes, so a model optimised under one solver can perform poorly when evaluated under another.

Table 3. Extrapolation error (mean MSE ± std) for different training (row) and testing (column) solvers.

Training	Testing			
	dopri5	euler	midpoint	rk4
dopri5	5.91e-02 ± 5.91e-02	5.25e-02 ± 2.72e-02	6.11e-02 ± 6.06e-02	5.91e-02 ± 5.91e-02
euler	1.02e-01 ± 4.87e-02	5.25e-02 ± 5.06e-02	1.04e-01 ± 4.95e-02	1.02e-01 ± 4.86e-02
midpoint	5.77e-02 ± 5.82e-02	5.24e-02 ± 2.60e-02	5.96e-02 ± 5.97e-02	5.77e-02 ± 5.82e-02
rk4	5.91e-02 ± 5.91e-02	5.25e-02 ± 2.72e-02	6.11e-02 ± 6.06e-02	5.91e-02 ± 5.91e-02

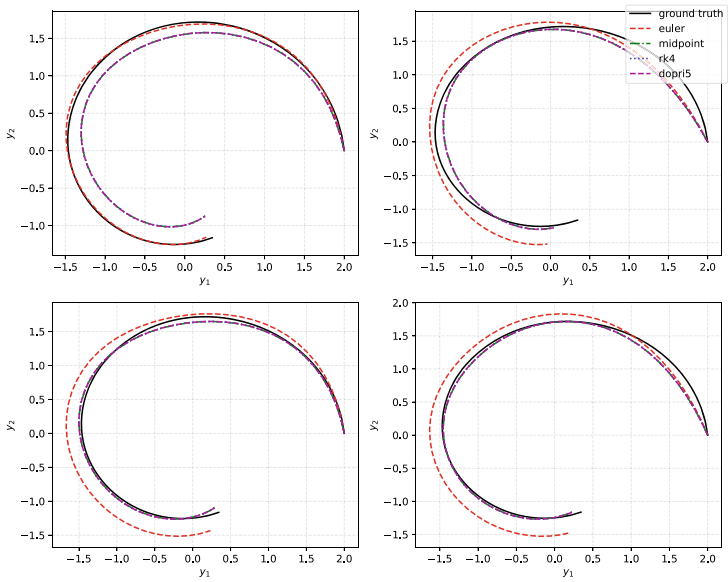


Fig. 1. Reconstruction test predictions for Neural ODE models trained and tested with *euler*, *dopri*, *rk4*, and *midpoint*. The same seed is used to allow to isolate the influence of the numerical method.

Table 3 and Fig. 2, show that errors grow substantially on the longer horizon, and here the differences between training solvers become more noticeable. Models trained with higher-order solvers (*dopri* or *midpoint*) that are then evaluated also with a high-accuracy solver achieve extrapolation MSE around $10e{-}2$, while training with *euler* leads to doubling the test error to $10e{-}1$. Note that, the low quality visual predictions in Fig. 2 are also caused by the extreme extrapolation test in which the models are predicting a time horizon twice as bigger than the one seen during training.

Overall, the numerical results support the statement that: (i) using a sufficiently accurate and consistent solver is important both for reconstruction and for long-term extrapolation; (ii) changing the solver between training and testing can have a large quantitative impact even when the training loss is nearly identical.

4.2 Effect of the Solver Order on the Optimisation Process

The previous experiments compare final errors across many solver configurations, but do not isolate how the solver affects the optimisation path itself. To do so, we perform a fixed-seed study that keeps all sources of randomness constant and varies only the training solver, and analyse the models' MSE over epochs for all solvers, Fig. 3.

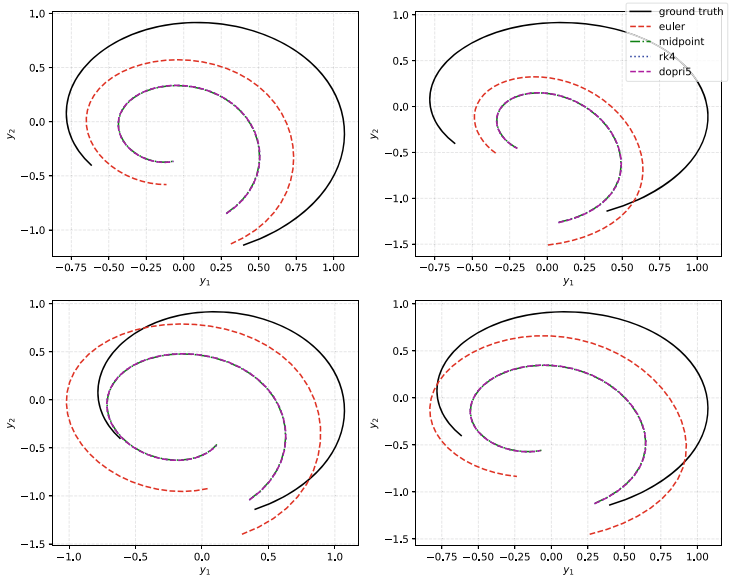


Fig. 2. Extrapolation test predictions for Neural ODE models trained and tested with *euler*, *dopri*, *rk4*, and *midpoint*. The same seed is used to allow to isolate the influence of the numerical method.

This experiment directly illustrates the gradient-level effects discussed in Sect. 3. Even when starting from the same initial parameters and observing the same data, different training solvers induce different effective gradients $\nabla \mathcal{L}^{\mathcal{A}_{\text{tr}}}(\boldsymbol{\theta})$, leading to different optimisation paths and converging to different parameter values and trajectories.

From Fig. 3, we observe that during the first 30 epochs all training solvers show almost identical MSE, consistent with the fact that the trajectories and hence the propagated numerical errors are still very small near the initialisation. As training progresses, solver-induced truncation error accumulates along the integration interval, leading to different effective gradients. In this regime *Euler* induces dynamics that decreases its effective loss faster and maintains the lowest training MSE during the whole training.

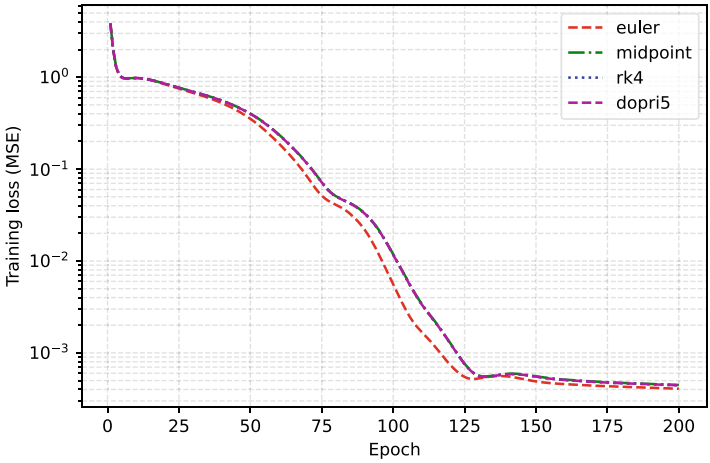


Fig. 3. Training MSE over epoch for four Neural ODE models trained from the same initialisation and data, but with different numerical integrators (*euler*, *dopri5*, *rk4*, *midpoint*).

5 Conclusion

In this work we demonstrated that the numerical integrator used in a Neural ODE is not a minor implementation detail, but an integral part of the model that shapes both the loss landscape and the optimisation dynamics. On the theoretical side, we proved that a fixed-step method of order p and step size h perturbs the ideal loss and its gradient by order $\mathcal{O}(h^p)$. Thus, different solvers induce different surrogate objectives and different gradient fields, with the discrepancy between two methods scaling with their respective global errors. Under some assumptions, we further demonstrated that the corresponding minimisers in the parameter space can be displaced from the ideal minimiser, and from each other, by h^p.

The numerical experiments on a simple spiral ODE corroborate these statement demonstrating that, in general, different numerical integrators can fit the training data to essentially the same MSE. Using different solvers in training and testing can increase prediction errors by up to an order of magnitude, even when the training loss is nearly identical. Specifically, mismatched train-test pairs involving a low-order solver, such as *euler*, produce larger discrepancies, while higher-order methods, when used consistently at train and test time, show markedly better performance Furthermore, starting from the same initial parameters and using the same data, models trained with different solvers initially follow similar loss curves, but their trajectories and parameter updates diverge as solver-induced truncation errors accumulate. This leads to distinct optimisation paths and different final parameter vectors, in line with our theoretical analysis of solver-dependent gradient perturbations.

As a whole, these results support the view that the choice of numerical method-order, and step size should be treated as part of the Neural ODE model specification. For reliable predictions, it is important to: (i) use a sufficiently accurate solver during both training and testing; (ii) avoid changing the solver (or its configurations) at prediction time without accounting for the induced bias; (iii) document the numerical configuration alongside architectural and optimisation details.

Beyond the specific results reported here, this work is as a call to action when developing, publishing, or deploying Neural ODEs and other solver-based neural networks, the numerical integrator and its settings should be treated as first-class modelling choices, documented with the same care as the architecture and optimisation procedure, and kept consistent between training and prediction whenever possible.

Note that, this theoretical analysis is valid to any neural network using a numerical method to output a response, such as Neural ODEs, and continuous-normalizing flows.

Future work could extend this analysis to stiff systems, adaptive-step training, and more complex architectures, as well as investigate principled strategies for jointly selecting or learning the numerical method and the Neural ODE parameters.

Acknowledgment. C. Coelho would like to thank the KIBIDZ project funded by dtec.bw - Digitalization and Technology Research Center of the Bundeswehr; dtec.bw is funded by the European Union-NextGenerationEU. We also acknowledge the computational resources (HPC cluster HSUper) that have been provided by the project hpc.bw, funded by dtec.bw.

References

1. Chapra, S.C.: Numerical Methods for Engineers. Mcgraw-Hill (2010)
2. Chen, R.T.Q.: torchdiffeq (2018). https://github.com/rtqichen/torchdiffeq
3. Chen, R.T., Rubanova, Y., Bettencourt, J., Duvenaud, D.K.: Neural ordinary differential equations. Adv. Neural Inf. Process. Syst. **31** (2018)
4. Coelho, C., Costa, M.F.P., Ferrás, L., Niggemann, O.: Neural ODEs for optimal control: learning continuous trajectories and feedback policies via adaptive discretisation. Authorea Preprints (2026)
5. Coelho, C., Costa, M.F.P., Ferrás, L.L.: Neural Chronos ODE: modelling bidirectional temporal patterns in time-series data. Exp. Syst. Appl. **273**, 126784 (2025)
6. Coelho, C., Costa, M.P., Ferrás, L.L.: Consistency matters: neural ODE parameters are dependent on the training numerical method. In: ICLR 2024 Workshop on AI4DifferentialEquations in Science (2024). https://openreview.net/forum?id=tPDA7Zt1bf
7. Coelho, C., Costa, M.F.P., Ferrás, L.L.: Neural fractional differential equations. Appl. Math. Model. **144**, 116060 (2025)
8. Dahlquist, G.: Convergence and stability in the numerical integration of ordinary differential equations. Mathematica Scandinavica, 33–53 (1956)

9. Hairer, E., Wanner, G., Nørsett, S.P.: Solving Ordinary Differential Equations I: Nonstiff Problems. SSCM. Springer, Heidelberg (1993). https://doi.org/10.1007/978-3-540-78862-1
10. Jia, J., Benson, A.R.: Neural jump stochastic differential equations. Adv. Neural Inf. Process. Syst. **32** (2019)
11. Khoo, Z.-Y., Lee, K.H., Huang, Z., Bressan, S.: Neural ordinary differential equations for the regression of macroeconomics data under the Green Solow model. In: Strauss, C., Kotsis, G., Tjoa, A.M., Khalil, I. (eds.) DEXA 2021. LNCS, vol. 12923, pp. 78–90. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-86472-9_7
12. Kumar, T., Kumar, A., Pal, P.: A physics-constrained neural ordinary differential equations approach for robust learning of stiff chemical kinetics. Combust. Theor. Model. **29**(3), 332–347 (2025)
13. Lee, K., Parish, E.J.: Parameterized neural ordinary differential equations: applications to computational physics problems. Proc. R. Soc. A Math. Phys. Eng. Sci. **477**(2253) (2021)
14. Li, J., Chen, W., Liu, Y., Yang, J., Zeng, D., Zhou, Z.: Neural ordinary differential equation networks for fintech applications using internet of things. IEEE Internet Things J. **11**(12), 21763–21772 (2024)
15. Nesterov, Y.: Introductory Lectures on Convex Programming Volume I: Basic Course. Lecture Notes, vol. 3, no. 4, p. 5 (1998)
16. Owoyele, O., Pal, P.: ChemNODE: a neural ordinary differential equations framework for efficient chemical kinetic solvers. Energy AI **7**, 100118 (2022)
17. Qian, Z., Zame, W., Fleuren, L., Elbers, P., van der Schaar, M.: Integrating Expert ODEs into Neural ODEs: pharmacology and disease progression. Adv. Neural. Inf. Process. Syst. **34**, 11364–11383 (2021)
18. Richter, T., Wang, W., Palma, A., Theis, F.J.: Generative models of cell dynamics: from Neural ODEs to flow matching. Commun. Biol. (2026)
19. Rubanova, Y., Chen, R.T., Duvenaud, D.K.: Latent ordinary differential equations for irregularly-sampled time series. Adv. Neural Inf. Process. Syst. **32** (2019)
20. Yang, Z., et al.: Detection of wind turbine blade abnormalities through a deep learning model integrating VAE and neural ODE. Ocean Eng. **302**, 117689 (2024)
21. Zakwan, M., Di Natale, L., Svetozarevic, B., Heer, P., Jones, C.N., Trecate, G.F.: Physically consistent neural odes for learning multi-physics systems. IFAC-PapersOnLine **56**(2), 5855–5860 (2023)
22. Zhu, Q., Guo, Y., Lin, W.: Neural delay differential equations. arXiv preprint arXiv:2102.10801 (2021)

Recurrent Neural Networks for Predicting Inflows in Hydroelectric Reservoirs

E. Villar[1], Cecília Coelho[2], M. Fernanda P. Costa[1(✉)], and L. L. Ferrás[3,4]

[1] Centre of Mathematics (CMAT), University of Minho, 4710-57 Braga, Portugal
mfc@math.uminho.pt
[2] Institute for Artificial Intelligence, Helmut Schmidt University, 22043 Hamburg, Germany
cecilia.coelho@hsu-hh.de
[3] Centro de Estudos de Fenómenos de Transporte (CEFT), Faculty of Engineering, University of Porto, 4200-465 Porto, Portugal
lferras@fe.up.pt
[4] ALiCE, Faculty of Engineering, University of Porto, 4200-465 Porto, Portugal

Abstract. Hydropower plays a crucial role in many countries' electricity systems, but its operation often conflicts with ecological preservation and agricultural water needs. Reservoirs must simultaneously deliver reliable energy, maintain minimum ecological flows to protect downstream ecosystems, and guarantee sufficient water for irrigation, making inflow forecasting critical for sustainable operation. In this work, we study neural network-based prediction of inflows to the Miranda hydropower reservoir in Brazil, aiming to provide accurate, data-driven forecasts that can support environmentally conscious and socially responsible decision-making. Using publicly available hydrological data, we perform a systematic exploratory analysis and pre-processing of hourly and daily time-series, including treatment of missing values, negative measurements, and outliers, as well as normalization and feature engineering. We then compare several recurrent neural network architectures (RNN, LSTM, BiLSTM, and GRU) under different training and model configurations. Our results show that a single-layer GRU model trained on the daily, normalized inflow series with a one-day lag achieves the best predictive performance. The selected model is used to produce 7-day ahead forecasts, which closely follow short-term dynamics while exposing limitations in anticipating abrupt peaks, thus highlighting the remaining challenges in forecasting extreme inflow events for hydropower management.

Keywords: Hydropower · Inflow Prediction · Neural Networks · Time-series · Reservoir Management · Data Analysis · Climate AI · Sustainability

1 Introduction

The construction of hydroelectric power plants and the consequent damming of rivers cause profound changes in the landscape and in environmental dynamics

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 272–284, 2027.
https://doi.org/10.1007/978-3-032-30530-5_17

[14]. These impacts are not limited to biological diversity, they also affect social and economic structures [2,7].

As an illustrative example, consider the basin of the Araguari River in Brazil. The sequence of dams built for hydroelectric power generation has led to deep transformations in the region. The formation of reservoirs has had a significant impact on communities that depend on fishing, both as a source of food and as a source of income. Their activities were practically extinguished, since most large fish species that were previously caught in the river cannot reproduce under the environmental conditions imposed by the reservoirs [12].

Another consequence of the construction of reservoirs is the difficulty in guaranteeing the multiple use of water resources. Historically, priority has been given to electricity generation over other potential uses, such as irrigation, recreation, and fish farming [16].

Hydroelectric projects also alter the ichthyological biodiversity of rivers, leading to changes in species composition and a reduction in species richness, which can in many cases result in the decline or even extinction of certain species [6,12].

Hydropower is one of the oldest and most widely used sources of electricity generation in the world and is often considered a relatively clean and renewable alternative compared with fossil fuels [11]. However, the development and operation of hydroelectric projects have sparked intense debate due to their environmental [1] and social [3] impacts.

As mentioned above, the absence of a minimum ecological discharge in hydroelectric projects can have devastating consequences for ecosystems and for local communities. Without an adequate release of water, downstream aquatic ecosystems suffer from a drastic reduction in water levels, which leads to the degradation of habitats that are essential for aquatic fauna and flora. This can result in a decline in biodiversity, affecting fish species and other forms of life that depend on a continuous flow of water to survive [15].

Another often neglected issue is the scarcity of water available to farmers. When water is retained in reservoirs for power generation, the downstream flow frequently decreases significantly, leaving farmers without sufficient access to the water resources needed to sustain their crops. This may cause crop losses, reduced agricultural productivity, and increased costs for farmers, who may be forced to seek alternative, often more expensive and less efficient, water sources [5].

Hydroelectric energy thus plays a crucial role in the provision of sustainable electricity worldwide. Nevertheless, the efficient and responsible operation of these plants poses substantial challenges, especially when it is necessary to balance the maximization of electricity production with ecosystem preservation, through compliance with prescribed ecological releases. In addition, hydroelectric plants must meet water allocation requirements for agricultural use, which further increases the complexity of water resources management [18].

The use of artificial intelligence (AI) in hydroelectric plant management has increased in recent years. In general, however, existing approaches focus primarily on maximizing energy production, without explicitly enforcing minimum

ecological discharges for ecosystem preservation and water allocation for agricultural irrigation [5,6].

Building on [5] and [6], this work focuses on the prediction component of a hybrid model that combines neural networks with constrained optimisation for the management of hydroelectric plants. The goal is to develop and analyse neural network models that can accurately predict reservoir inflows, providing reliable inputs to downstream optimisation procedures and contributing to more transparent and ethically grounded AI-based decision support systems. The goal of the predictive component is to forecast the volume of water that enters the hydroelectric reservoir at each time step. A neural network is implemented to predict these inflows, both on an hourly and on a daily scale. Accurate inflow prediction is crucial for any subsequent optimisation model, since it provides information about future water availability and thus enables more efficient management of current and future water resources.

This paper is organised as follows. Section 2 introduces the case study, the Miranda hydroelectric plant, describing its main characteristics and presenting an exploratory analysis and pre-processing of the hydrological dataset of reservoir inflows. Section 3 presents the neural network-based inflow prediction methodology and reports performance results along with a comparison discussion. Finally, Sect. 4 concludes the work and discusses future research directions.

2 Case Study: Miranda Hydroelectric Plant

In this work, the Miranda hydroelectric plant is considered as a case study since its hydrological data are publicly available in [13] (a database provided in Portuguese), with hourly and daily measurements. The plant is located on the Araguari River, in the municipality of Indianópolis, Minas Gerais, Brazil (see Fig. 1). The reservoir covers an area of 50.61 km^2 and has a useful storage volume of $1.12 \times 10^9 m^3$. The dam has a maximum height of 79 m and a length of 2050 m, with an effective head of 66.4 m for power generation. The installed capacity is 408 MW, distributed across three generating units equipped with vertical Francis turbines of 136 MW each. The firm energy for commercialisation is, on average, 198.2 MW [8].

In this study, the variables of interest are "val_vazaoafluente" and "din_instante". This dataset will be referred to simply as the *dataset*. As mentioned before, measurements are available at an hourly resolution. Time stamps specify year, month, day, and hour, for example: "2021-01-01 01:00:00", "2021-01-01 02:00:00". Figure 2 shows the inflows over time (variable "val_vazaoafluente") for the *dataset* (Table 1).

2.1 Exploratory Data Analysis

Data quality analysis. We start by analysing the variable "val_vazaoafluente". This variable contains one missing value (NaN) and 1434 negative values. Since the physical meaning of this variable is not compatible with negative values, as

Fig. 1. Location of the Miranda hydroelectric plant.

Table 1. Description of variables in the hydrological dataset

Attribute	Type	Description/Values
id_subsistema	Qualitative (nominal)	Subsystem identifier: {SE, S, N, NE}
nom_subsistema	Qualitative (nominal)	Basin regions: {Sudeste/Centro-Oeste, Sul, Norte, Nordeste}
tip_reservatório	Qualitative (nominal)	Reservoir type: Fio d'água; Reservatório com centrais; Reservatório sem centrais; Reservatório com centrais de bombeamento
nom_bacia	Qualitative (nominal)	Basin name (e.g., Amazonas, Uruguai)
id_reservatorio	Qualitative (nominal)	Reservoir identifier (e.g., AMUHBB)
nom_reservatorio	Qualitative (nominal)	Reservoir name (e.g., Balbina, São Roque)
cod_central	Quantitative (discrete)	Power plant code (e.g., 88, 277)
din_instante	Qualitative (ordinal)	Timestamp (year, month, day, hour)
val_vazaoafluente	Quantitative (continuous)	Inflow discharge to the reservoir
val_nivelmontante	Quantitative (continuous)	Upstream water level
val_niveljusante	Quantitative (continuous)	Downstream water level
val_vazaodefluente	Quantitative (continuous)	Outflow discharge from the reservoir
val_vazaoturbinada	Quantitative (continuous)	Turbined flow used for energy generation
val_vazaovertida	Quantitative (continuous)	Spilled flow not used for energy generation
val_vazaotransferida	Quantitative (continuous)	Water transferred to another reservoir/location
val_vazaovertidanaoturbinavel	Quantitative (continuous)	Spilled flow not passing through turbines and not usable for generation

Dataset composed of 26,274 records and 18 attributes, constructed from hourly hydrological data files covering the period 2021–2023 [13].

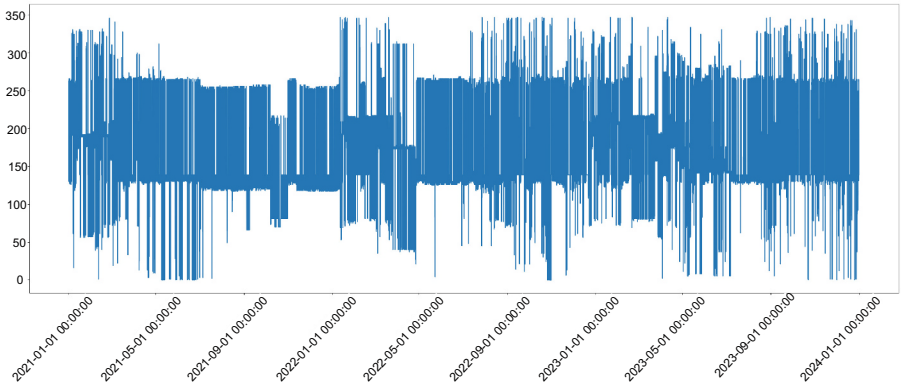


Fig. 2. Visualization of the inflows over time.

it represents the volume of water entering the reservoir, these negative entries were treated as measurement errors and thus removed. Furthermore, Fig. 3 shows a highly right-skewed distribution for this variable in which most observations are concentrated in a relatively low range, with the median located well below the upper whisker. The interquartile range (IQR) measure is fairly compact, indicating that typical inflow values do not vary dramatically. However, there is a large number of outliers (with 3473 identified) extending far above the upper whisker, reaching values above 1200. This suggests the presence of extreme inflow events which can be associated with heavy rainfall periods or seasonal hydrological peaks. The lower whisker is close to zero, showing that very low inflow conditions also happen, but without significant lower-end outliers.

Data Cleaning. Guided by the data analysis, a copy of the dataset was created in which outliers (the NaN value, and the negative values) were replaced by the median. These replacements affect approximately 19% of the data.

Other strategies were also tested: replacing negative values by zero; replacing outliers with the mean of their adjacent neighbours; and using the *SimpleImputer* class from *sklearn.impute* with the strategies *mean*, *median*, *most_frequent*, and *constant*. It was observed that replacing these values by the median leads to better performance when modelling inflows, measured in terms of the mean squared error (MSE).

Table 2 reports the four best MSE values obtained with these strategies, denoted E1 to E4.

Thus, replacing outliers and negative values by the median yields the lowest MSE. The other imputation strategies led to higher MSE values than those reported in Table 2. Since strategy E4 achieves the lowest MSE, this cleaned version of the data is used in the subsequent analysis and will be referred to as Dataset-H.

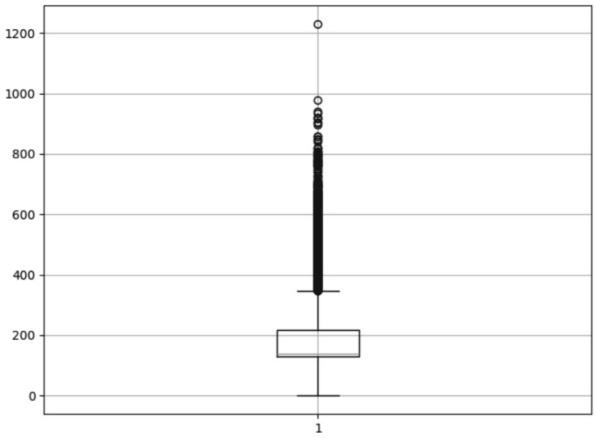


Fig. 3. Boxplot of the inflow discharge variable "val_vazaoafluente". The distribution is strongly right-skewed, with a compact interquartile range and numerous high-value outliers.

Table 2. Best MSE values obtained for different strategies for replacing outliers and negative values.

Strategy	Negative values	Outliers	MSE
E1	Neighbour mean	Neighbour mean	5337
E2	Neighbour mean	Not replaced	10545
E3	0	Median	3354
E4	Median	Median	2628

The attribute "din_instante" was converted to *datetime* format, ensuring that the data follow a correct chronological order.

For the purposes of this work, a daily dataset was also created to reduce the temporal resolution. To this end, hourly inflows were aggregated by day using a sum, yielding a daily dataset, Dataset-D, derived from Dataset-H.

Data Analysis. We first investigate how many past lags (days or hours) are most appropriate for predicting the inflow for the next day (or hour).

As an initial step, we study the autocorrelation function, which measures the correlation of a time-series with itself at different lags. Figure 4 shows the autocorrelation function (ACF) for Dataset-D at different time lags, helping to identify repeating patterns (seasonality). The results indicate that autocorrelation is strongest at with a time lag of 1, suggesting that inflows are most strongly correlated with the previous day.

The partial autocorrelation function (PACF) measures the correlation between a time-series and its own lags, while controlling for the effect of intermediate lags. While the ACF captures both direct and indirect correlations between

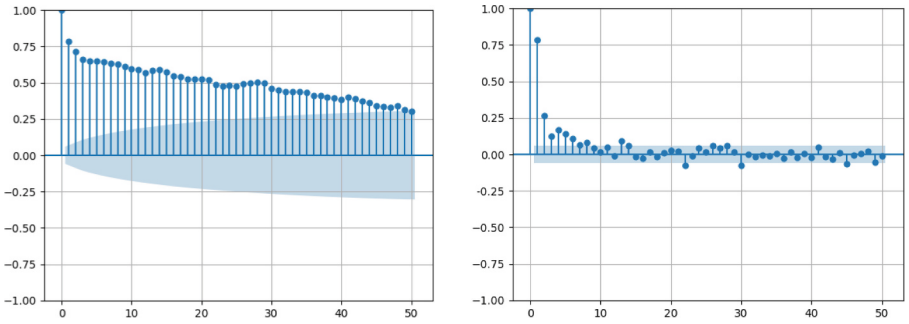


Fig. 4. Autocorrelation (left) and partial autocorrelation (right) function over different time lags for Dataset-D.

current and past values, the PACF removes the effect of preceding lags and focuses on the direct correlation. Figure 4 shows the PACF for different lags. The PACF is useful for identifying the order of autoregressive models, where significant peaks indicate lags that have a direct relationship with the current value after removing the influence of earlier lags. As with the ACF, the largest partial autocorrelation occurs at a time lag of 1.

Note that, the results for Dataset-H are not displayed due to the similarity with the ones obtained for Dataset-D.

From this analysis, only the previous day (or hour) is used in the pre-processing stage to predict the inflow at the current day (or hour).

Since Dataset-D has a coarser temporal resolution, an additional study was performed to decompose the data into trend, seasonality, and residual components. For this purpose, the `seasonal_decompose` function from the *statsmodels* Python package was used. This decomposition is useful for understanding the dynamics of inflows over time, separating long-term trends, seasonal behaviour, and unexplained fluctuations. In the pre-processing stage we will explore whether including trend, seasonal, and residual components as additional features benefits the predictive performance.

Note that the attributes used as inputs correspond to the previous day, and the inflow to be predicted for the current day is not fed back as an input.

Figure 5(a) shows the daily inflow time-series. The series exhibits high variability, with significant fluctuations. Some seasonal oscillations can be observed, with inflow peaks in certain periods followed by sharp drops. There appears to be an overall stabilization or slight decline in the most recent data, after a period of increase at the beginning of 2022.

In Fig. 5(b), the trend reflects the long-term direction of the data, smoothing short-term fluctuations. It suggests a gradual increase in inflows from mid-2021 until the end of 2022. From 2023 onwards, there is a slight deceleration or decline in the trend, indicating stabilization or a potential decrease in reservoir inflows.

Figure 5(c) shows the seasonal component, which represents repetitive and predictable patterns occurring at regular intervals. It reveals marked seasonal

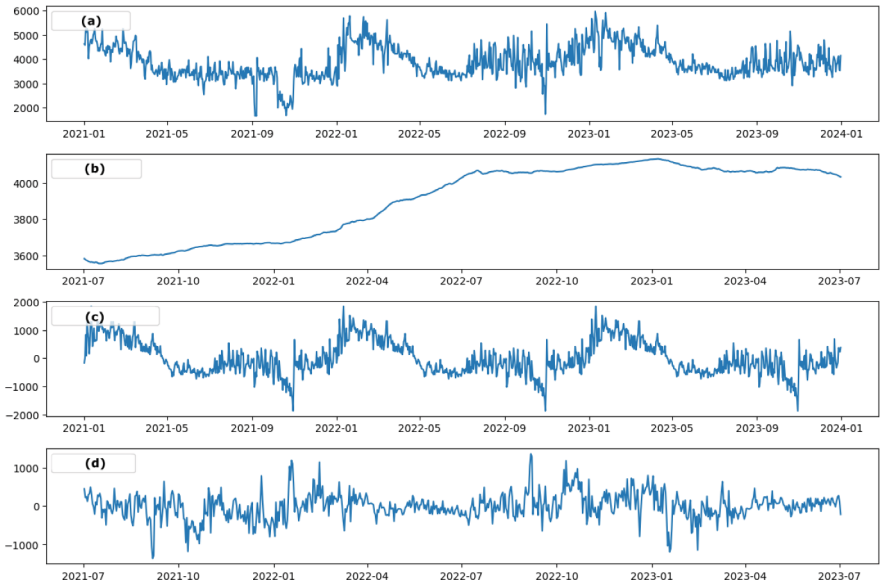


Fig. 5. Data decomposition analysis for Dataset-D: (a) original daily inflow series; (b) trend; (c) seasonal component; (d) residuals.

variations, with peaks and troughs that likely correspond to climatic cycles such as rainy and dry seasons.

Figure 5(d) displays the residual component, which contains the remaining variability after removing trend and seasonality from the original series. The residuals appear largely random, with no clear pattern, indicating that most of the structure in the series has been captured by the trend and seasonal components. A few pronounced peaks remain, which may correspond to isolated events or anomalies not explained by the other components.

Pre-processing. As discussed above, we use a time lag of 1, *i.e.* one day for Dataset-D and one hour for Dataset-H, and the target variable $\{y_t\}_{t=1}^{T}$ is the inflow volume to the reservoir at a given day (or hour). Several pre-processing techniques were investigated to improve the performance of the neural network models, including log-transformation, differencing, standardisation, noise injection, normalisation, and the creation of additional features:

- Log transformation: Applying a logarithmic transformation to the target using *numpy*, i.e., $y \leftarrow \ln(y_t)$, where ln denotes the natural logarithm;
- Differencing: A standard technique in time-series analysis to obtain stationarity by removing long-term trends or seasonal patterns. The first difference is computed as $d_t = y_t - y_{t-1}$;
- Standardization: Transforming the target to have zero mean and unit variance, $y_t \leftarrow \dfrac{y_t - \mu}{\sigma}$, where μ and σ denote the mean and standard deviation

of y, respectively. This places all variables on a common scale so that they contribute proportionally during model training;
- Noise injection: Adding noise to the training data, in which the noise level controls the amount of random perturbation added to the data and is often defined as a fraction of the standard deviation, since σ reflects the natural scale of the data. Noise injection can increase model robustness and help prevent overfitting by encouraging the model to tolerate small variations: $y_t \leftarrow y_t + \varepsilon$, where ε is a random perturbation with prescribed variance;
- Normalisation: Scaling the data to a fixed interval. Here, values are mapped to $[0, 1]$ via $y_t \leftarrow \dfrac{y_t - y_{\min}}{y_{\max} - y_{\min}}$, where $y_{\min}$ and $y_{\max}$ are the minimum and maximum observed values of $(y_0, y_1, \ldots, y_N)$. This is particularly useful since neural networks are sensitive to the magnitude of input variables;
- Feature engineering: Using the `seasonal_decompose` function from *statsmodels* to decompose the target into trend, seasonal, and residual components and including these as additional features. For Dataset-H, additional calendar features were also created from "din_instante": "day", "month", and "hour". Using these features and the elbow method, a clustering step with 11 clusters was performed, resulting in an additional categorical feature "cluster".

Among all tested techniques, normalisation yielded the best predictive performance in terms of the adopted evaluation metrics. Therefore, normalisation was chosen as the main pre-processing step for the development of this work.

3 Neural Networks for Inflow Prediction

Different recurrent neural network architectures were tested: Long Short-Term Memory (LSTM) [9], Simple Recurrent Neural Network (RNN) [10], Bidirectional LSTM (BiLSTM) [17], and Gated Recurrent Unit (GRU) [4]. In addition, several network design choices were explored, including increasing or decreasing the number of neurons per layer, changing the number of layers, adding regularization and/or dropout layers, modifying the learning rate and/or the optimiser, changing the activation functions, and varying the weight initialisation schemes.

Both datasets were split into 80% training, 10% test, and 10% validation sets, while preserving the chronological order.

Note that, the neural network architectures were tested on both hourly (Dataset-H) and daily (Dataset-D) datasets. However, while the finer temporal granularity of Dataset-H captures short-term fluctuations, it also amplifies noise from measurement errors and irregular hourly variations, leading to unstable training dynamics and inferior generalisation performance across all architectures. For these reasons, we focus subsequent analysis and modelling exclusively on Dataset-D.

Table 3 reports the root mean squared error (RMSE) and mean absolute percentage error (MAPE) for the different strategies and approaches, using Dataset-D. The acronym NSOVN indicates there was no replacement of outliers and negative values.

Table 3. Results for different strategies and networks for the daily data, Dataset-D.

Scenario	Strategy	Method	Network	RMSE	MAPE
1	E4	Normalisation	RNN	508.6	11
2	NSOVN	Normalisation	RNN	606.6	16
3	E4	Normalisation	LSTM	611.8	14
4	E4	Normalisation	GRU	364.9	7

All networks were trained with the Adam optimiser for 100 epochs, using a batch size of 64 and an early-stopping callback with patience equal to 10. The models are configured to generate forecasts over a 7-day horizon in a recursive manner, *i.e.*, the predicted value for a given day is fed back as input to predict the inflow for the next day. All implementations were done using *Pytorch*, while data analysis and preprocessing was done using *Pandas* and *Scikit-learn*.

The analysis of Table 3 shows that Scenario 4 achieves the best performance. The selected model consists of a single GRU layer with 100 neurons, taking as input the inflow value of the previous day, followed by a dropout layer with dropout rate 0.1. The training process of the chosen network was analysed by monitoring the evolution of the training loss and validation loss (val_loss), both computed as MSE, Fig. 6. The rapid decrease of the loss during the first epochs indicates that the model is able to learn effectively from the data, with only minor improvements after approximately four epochs. Moreover, the training and validation loss curves remain close and do not diverge significantly, suggesting that the model is well fitted and does not exhibit clear signs of overfitting.

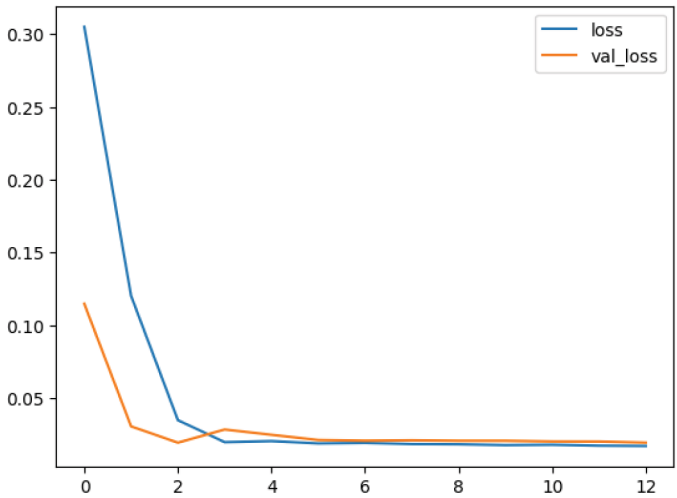


Fig. 6. Evolution of the training and validation loss for the best performant model.

The early-stopping callback interrupted training after 12 epochs since the validation performance did not improve for 10 consecutive epochs. Other configurations, such as different weight initialisations, additional dropout layers, regularisation, alternative activation functions, or modified learning rates did not lead to further performance improvements.

Table 4 reports the predicted inflows for the 7-day horizon, together with the corresponding observed values.

Table 4. Predicted versus observed inflow values for a 7-day horizon.

Day	Observed value	Prediction
1	3581	3596
2	3598	3513
3	3522	3471
4	3516	3439
5	4074	3414
6	3932	3396
7	3828	3382

From Table 4, it can be seen that, for the first three days, the model performance is relatively good as the forecasts follow the observed trend, and the error remains within an acceptable range. On day 4, however, a significant event (a pronounced peak) occurs in the observed series, which the model fails to predict. Due to the recursive forecasting scheme, any prediction error is fed back as if it were a true input, and thus tends to accumulate and amplify over time. Another contributing factor may be the limited ability of the network to capture abrupt changes as the model appears to have learnt a smoothed pattern, and consequently fails to reproduce this anomalous event. Despite these challenges, the overall results confirm the effectiveness of neural networks, in particular GRU, for reliable short-term inflow prediction in hydropower applications.

4 Conclusion

This work has demonstrated the feasibility of neural network-based inflow prediction for the Miranda hydroelectric reservoir using three years of hydrological data. Through exploratory data analysis, cleaning strategies, and feature engineering, we constructed a high-quality hourly (Dataset-H) and daily (Dataset-D) time-series, with the daily dataset proving superior for predictive modelling.

Among the tested recurrent architectures (Recurrent Neural Network, Long-Short Term Memory, Bidirectional Long-Short Term Memory, and Gated Recurrent Unit), a single-layer Gated-Recurrent Unit model with 100 neurons, trained on normalised daily inflows using a one-day lag, achieved the best performance, with root mean square error of 364.9 and mean absolute percentage error of

7%. The 7-day ahead forecasts accurately captured short-term dynamics but struggled with abrupt peaks, demonstrating the challenges of modelling extreme inflow events in recursive schemes where errors accumulate over time.

These findings provide a reliable predictive component for sustainable reservoir management, where accurate inflow forecasts are essential to balance energy production with ecological minimum flows and agricultural water demands.

Future work will focus on improving peak prediction through hybrid architectures (e.g., attention mechanisms or CNN-LSTM ensembles), incorporating meteorological covariates, and extending the forecast horizon while maintaining low error rates. Integrating this predictor into a constrained optimisation framework will enable sustainable operating policies for the Miranda plant.

Acknowledgments. We thank the support of the High-Performance Computing Center at the University of Évora funded by FCT I.P. under the project "OptXAI: Constrained optimisation in NNs for Explainable, Ethical and Greener AI", reference 2024.00191.CPCA.A1, platform Vision. This work was financially supported by Fundação para a Ciência e Tecnologia (FCT) through the FCT/MCTES (PIDDAC) under the project 2022.06672.PTDC-iMAD (Improving the Modelling of Anomalous Diffusion and Viscoelasticity: solutions to industrial problems), DOI 10.54499/2022.06672.PTDC (https://doi.org/10.54499/2022.06672.PTDC); and by the projects LA/P/0045/2020 (ALiCE), UIDB/00532/2020, and UIDP/00532/2020 (CEFT), and UID/00013/2025 (https://doi.org/10.54499/UID/00013/2025) (CMAT). It was also financially supported by Fundação "la Caixa"|BPI and FCT through project PL24-00057: "Inteligência Artificial na Otimização da Rega para Olivais Resilientes às Alterações Climáticas". C. Coelho would like to thank the KIBIDZ project funded by dtec.bw-Digitalization and Technology Research Center of the Bundeswehr; dtec.bw is funded by the European Union-NextGenerationEU.

References

1. Botelho, A., Ferreira, P., Lima, F., Pinto, L.M.C., Sousa, S.: Assessment of the environmental impacts associated with hydropower. Renew. Sustain. Energy Rev. **70**, 896–904 (2017)
2. Catolico, A., Maestrini, M., Strauch, J., Giusti, F., Hunt, J.: Socioeconomic impacts of large hydroelectric power plants in Brazil: a synthetic control assessment of Estreito hydropower plant. Renew. Sustain. Energy Rev. **151**, 111508 (2021)
3. Cernea, M.M., et al.: Social impacts and social risks in hydropower programs: preemptive planning and counter-risk measures. In: Keynote address: Session on social aspects of hydropower development. United Nations Symposium on Hydropower and Sustainable Development Beijing, China (2004)
4. Cho, K.: Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078 (2014)
5. Coelho, C., Costa, M.F.P., Ferrás, L.L.: The influence of neural networks on hydropower plant management in agriculture: addressing challenges and exploring untapped opportunities. In: Mercier-Laurent, E., Kayakutlu, G., Owoc, M.L., Wahid, A., Mason, K. (eds.) IFIP International Workshop on Artificial Intelligence for Knowledge Management, pp. 150–160. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-61069-1_11

6. Coelho, C., Jing, M., Costa, M.F.P., Ferrás, L.L.: An adaptive hydropower management approach for downstream ecosystem preservation. arXiv preprint arXiv:2403.02821 (2024)
7. De Faria, F.A., Davis, A., Severnini, E., Jaramillo, P.: The local socio-economic impacts of large hydropower plant development in a developing country. Energy Econ. **67**, 533–544 (2017)
8. ENGIE: Miranda hydroelectric power plant (2026). https://www.engie.com.br/en/usinas/miranda-hydroelectric-power-plant/. Accessed 26 Feb 2026
9. Hochreiter, S.: Long short-term memory. Neural Comput. (1997)
10. Keras: SimpleRNN. https://keras.io/api/layers/recurrent_layers/simple_rnn/. Accessed 26 Feb 2026
11. Killingtveit, Å.: Hydropower. In: Managing Global Warming, pp. 265–315. Elsevier (2019)
12. Loures, R.C., Pompeu, P.S.: Temporal changes in fish diversity in lotic and lentic environments along a reservoir cascade. Freshw. Biol. **64**(10), 1806–1820 (2019)
13. ONS: Dados hidrologicos horario 2024-01 (2024). https://dados.ons.org.br/dataset/dados_hidrologicos_ho/resource/ca306dc0-abbe-47d2-af9d-da601fc2c09f?inner_span=True. Accessed 26 Feb 2026
14. Pandit, M.K., Grumbine, R.E.: Potential effects of ongoing and proposed hydropower development on terrestrial biological diversity in the Indian Himalaya. Conserv. Biol. **26**(6), 1061–1071 (2012)
15. Renöfält, B.M., Jansson, R., Nilsson, C.: Effects of hydropower generation and opportunities for environmental flow management in Swedish riverine ecosystems. Freshw. Biol. **55**(1), 49–67 (2010)
16. Ribas, J.R.: An assessment of conflicting intentions in the use of multipurpose water reservoirs. Water Resour. Manage. **28**(12), 3989–4000 (2014)
17. Schuster, M., Paliwal, K.K.: Bidirectional recurrent neural networks. IEEE Trans. Sig. Process. **45**(11), 2673–2681 (1997)
18. Zeng, R., Cai, X., Ringler, C., Zhu, T.: Hydropower versus irrigation—an analysis of global patterns. Environ. Res. Lett. **12**(3), 034006 (2017)

Parameter Estimation for a Springpot–Fractional Dashpot Viscoelastic Model: Optimisation and Numerical Laplace Inversion

L. L. Ferrás[1,2(✉)], C. Coelho[3], M. Fernanda P. Costa[2], M. L. Morgado[4,5], and M. Rebelo[6]

[1] CEFT, ALiCE, Faculty of Engineering, University of Porto, 4200-465 Porto, Portugal
lferras@fe.up.pt

[2] Centre of Mathematics (CMAT), University of Minho, 4710-57 Braga, Portugal
mfc@math.uminho.pt

[3] Institute for Artificial Intelligence, Helmut Schmidt University, 22043 Hamburg, Germany
cecilia.coelho@hsu-hh.de

[4] Departamento de Matemática, Universidade de Trás-os-Montes e Alto Douro, UTAD, 5001-801 Vila Real, Portugal
luisam@utad.pt

[5] Center for Computational and Stochastic Mathematics, Instituto Superior Técnico, Lisbon, Portugal

[6] Center for Mathematics and Applications (NOVA Math) and Department of Mathematics, NOVA School of Science and Technology (NOVA FCT), Campus de Caparica, Caparica2829-516, Portugal
msjr@fct.unl.pt

Abstract. This work addresses the parameter estimation of a fractional viscoelastic model comprising a springpot and a fractional dashpot arranged in series. Key features of the model, including the relaxation modulus, are formulated in the Laplace domain and subsequently evaluated in the time domain through numerical inversion using the Talbot method.

A two-step optimisation strategy is adopted to identify the model parameters efficiently. In the first stage, frequency-domain data, namely the storage and loss moduli, are used to estimate an initial set of parameters. In the second stage, these preliminary estimates are refined by performing an optimisation that incorporates both frequency-domain data and time-domain data (relaxation modulus).

Since the relaxation modulus is obtained through an inverse Laplace transform and does not admit a closed-form expression in the time domain, a sequential iterative procedure is employed during the second stage. In this approach, the relaxation modulus is first computed numerically using the Talbot method, after which the parameters are updated through optimisation. This process is repeated iteratively until a prescribed tolerance is satisfied.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 285–297, 2027.
https://doi.org/10.1007/978-3-032-30530-5_18

Validation against experimental data obtained for a low-density polyethylene demonstrates that the proposed methodology accurately reproduces both frequency- and time-domain responses, highlighting the effectiveness of fractional models in capturing complex viscoelastic behaviour.

Keywords: Optimisation · Viscoelastic models · Rheology · Talbot Method · Springpot · Fractional-Dashpot

1 Introduction

The modelling of viscoelastic fluids [1] using fractional elements [3–7,10,15] has proven to be an effective way to capture complex relaxation behaviour with a reduced number of parameters. In this context, the fractional dashpot model introduced by Yao [20] provides a continuous interpolation between viscous and Maxwell-type responses, while retaining a compact mathematical structure. Nevertheless, many complex fluids exhibit a distribution of relaxation mechanisms that cannot be adequately described by a single fractional dashpot and inherent timescales.

To overcome this limitation, we propose an extended configuration consisting of a springpot connected in series with a fractional dashpot [8]. This construction preserves the dissipative nature of the original formulation, while introducing additional flexibility to represent a broader class of material responses. In contrast with classical multimode approaches, which rely on a superposition of many discrete relaxation modes, the present formulation incorporates multiple memory effects through fractional operators, leading to a smoother and more compact representation of the relaxation spectrum.

From a physical perspective, the series arrangement can be interpreted as the interaction of two distinct dissipative mechanisms acting at different temporal scales. This interpretation is consistent with the hierarchical relaxation processes typically observed in polymeric and complex fluids. At the same time, the use

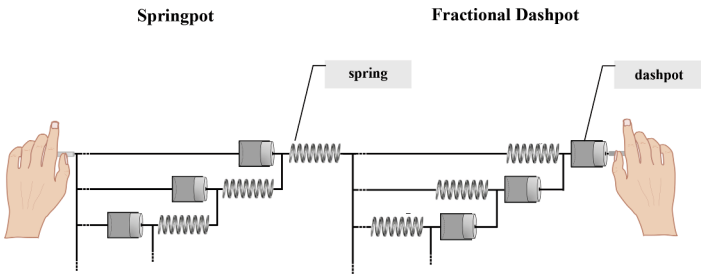


Fig. 1. Mechanical analogue consisting of a springpot and a fractional dashpot connected in series. The ladder-type representation associated with the fractional dashpot follows the construction introduced in Yao's model.

of fractional operators ensures a parsimonious description, avoiding the need for a large number of fitting parameters.

We consider a mechanical system composed of a springpot and a fractional dashpot connected in series, as depicted in Fig. 1.

The strain in the springpot element is given by

$$\gamma_1(t) = \frac{1}{\mathbb{V}} D_t^{-\alpha} \sigma(t),$$

where $D_t^{-\alpha}$ denotes the Riemann-Liouville fractional integral [2] of order $\alpha \in (0,1)$.

For the fractional dashpot, the strain rate is expressed as

$$\dot{\gamma}_2(t) = \frac{e^{-t/\lambda}}{g\lambda^\beta} D_t^{1-\beta} \left(e^{t/\lambda} \sigma(t) \right),$$

with $\beta \in (0,1)$, where $D_t^{1-\beta}$ denotes the Caputo fractional derivative of order $(1-\beta)$ [2].

Since the elements are arranged in series, the total strain satisfies

$$\gamma(t) = \gamma_1(t) + \gamma_2(t).$$

Combining the previous expressions leads to the constitutive relation [8],

$$g\lambda^\beta D_t^{1-\alpha} \sigma(t) + \mathbb{V} e^{-t/\lambda} D_t^{1-\beta} \left(e^{t/\lambda} \sigma(t) \right) = \mathbb{V} g \lambda^\beta \dot{\gamma}(t). \tag{1}$$

Here, g, $\mathbb{V}$ and λ are positive material parameters. D_t^μ denotes the Caputo fractional derivative of order $\mu \in (0,1)$, defined for a sufficiently smooth function f by,

$$D_t^\mu f(t) = \frac{1}{\Gamma(1-\mu)} \int_0^t (t-\tau)^{-\mu} f'(\tau)\, d\tau, \quad 0 < \mu < 1.$$

The original fractional dashpot model proposed by Yao is recovered as a limiting case of (1) by dividing through by $\mathbb{V}$ and considering the limit $\mathbb{V} \to \infty$.

This work is organised as follows. Section 2 introduces the relevant rheological functions, including the storage and loss moduli and the relaxation modulus. In Sect. 3, the Talbot method is presented, together with the optimisation problems used to estimate the model parameters. Section 4 describes the optimisation strategy adopted to solve these problems, and demonstrates that the optimal parameter values enable the model to reproduce the experimental data with high accuracy while using a smaller number of parameters. Finally, the paper concludes in Sect. 5.

2 Rheological Functions

In this section, we present key rheological functions associated with the constitutive model (1), namely the storage modulus, the loss modulus, and the relaxation

modulus. These functions provide essential information about the material's viscoelastic behaviour in both the frequency and time domains.

For parameter estimation, only these three functions are considered. The storage and loss moduli, measured in the frequency domain, are used to characterise the oscillatory response of the material, while the relaxation modulus, measured in the time domain, captures the stress relaxation behaviour. Together, they provide sufficient and complementary information to reliably identify the model parameters and demonstrate the robustness of the proposed fitting methodology.

2.1 Storage and Loss Modulus

The storage modulus $G'(\omega)$ and the loss modulus $G''(\omega)$ are obtained from the complex modulus $G^*(\omega)$ by evaluating the Laplace-domain transfer function along the imaginary axis, i.e., $s = i\omega$. They are given by [8]:

$$G'(\omega) = \frac{\mathbb{V} g \lambda^\beta \, \omega \, B(\omega)}{A^2(\omega) + B^2(\omega)},$$
$$G''(\omega) = \frac{\mathbb{V} g \lambda^\beta \, \omega \, A(\omega)}{A^2(\omega) + B^2(\omega)},$$

where the auxiliary functions $A(\omega)$ and $B(\omega)$ correspond to the real and imaginary parts of the denominator, respectively, and are defined as

$$A(\omega) = g\lambda^\beta \omega^{1-\alpha} \cos\left(\frac{(1-\alpha)\pi}{2}\right) + \mathbb{V}\left(\omega^2 + \lambda^{-2}\right)^{\frac{1-\beta}{2}} \cos\big((1-\beta)\phi(\omega)\big),$$
$$B(\omega) = g\lambda^\beta \omega^{1-\alpha} \sin\left(\frac{(1-\alpha)\pi}{2}\right) + \mathbb{V}\left(\omega^2 + \lambda^{-2}\right)^{\frac{1-\beta}{2}} \sin\big((1-\beta)\phi(\omega)\big),$$

with

$$\phi(\omega) = \arctan(\omega\lambda).$$

2.2 Relaxation Modulus

The relaxation modulus $G(t)$ is defined through the stress response to a unit step strain. In the Laplace domain, it is given by [8]:

$$\tilde{G}(s) = \frac{\mathbb{V}}{s^{1-\alpha} + a(s+b)^{1-\beta}},$$

where

$$a = \frac{\mathbb{V}}{g\lambda^\beta}, \qquad b = \lambda^{-1},$$

and $s \in \mathbb{C}$ denotes the Laplace variable.

The time-domain relaxation modulus is then obtained via the inverse Laplace transform:

$$G(t) = \mathbb{V}\,\mathcal{L}^{-1}\left\{\frac{1}{s^{1-\alpha} + \dfrac{\mathbb{V}}{g\lambda^{\beta}}\,(s+\lambda^{-1})^{1-\beta}}\right\}(t).$$

3 Optimisation Problems and Talbot Method

3.1 Numerical Inversion via the Talbot Method

Time-domain rheological functions, such as the relaxation modulus $G(t)$, are obtained as inverse Laplace transforms of their corresponding Laplace-domain expressions. In fractional viscoelastic models, these Laplace-domain expressions typically involve fractional powers of the Laplace variable and special functions, for which closed-form inverse transforms are rarely available. Consequently, reliable and efficient numerical inversion techniques are required to evaluate the time-domain responses.

The inverse Laplace transform of a function $\tilde{f}(s)$ is formally defined by the Bromwich integral

$$f(t) = \frac{1}{2\pi i}\int_{\mathbb{L}} e^{st}\tilde{f}(s)\,ds, \qquad t > 0,$$

where $\mathbb{L}$ denotes a vertical contour in the complex plane, defined by

$$\mathbb{L} = \{s = \gamma + i\omega : \omega \in (-\infty, \infty)\},$$

with γ chosen such that all singularities of $\tilde{f}(s)$ lie to the left of the contour. Although this representation is mathematically well-defined, its direct numerical evaluation is challenging due to the oscillatory behaviour of the exponential term e^{st}, especially for large values of $|\omega|$. Additionally, the integrand often decays slowly, making standard quadrature methods inefficient and potentially unstable.

To overcome these difficulties, Talbot proposed deforming the Bromwich contour into a more suitable path in the complex plane [13]. The resulting contour, commonly referred to as the Talbot contour, is designed to ensure rapid decay of the exponential term while maintaining analyticity of the integrand. This contour deformation relies on Cauchy's integral theorem, which guarantees that the value of the inverse Laplace transform remains unchanged as long as the deformed contour does not cross any singularities of $\tilde{f}(s)$.

The Talbot contour is typically parameterized as

$$s(\theta) = \frac{N}{t}\,\phi(\theta), \qquad -\pi \le \theta \le \pi,$$

where $\phi(\theta)$ is a suitably chosen complex-valued function that defines the contour shape, and N is the number of quadrature points. A commonly used parameterization is given by

$$\phi(\theta) = a\,\theta\cot(\theta) + ib\,\theta + c,$$

where a, b, and c are real parameters selected to optimise convergence and numerical stability. This contour bends into the left half-plane, ensuring that the real part of $s(\theta)$ is negative for most values of θ, thereby enforcing exponential decay in e^{st}.

After parameterization, the Bromwich integral is transformed into a contour integral in terms of the parameter θ,

$$f(t) = \frac{1}{2\pi i} \int_{-\pi}^{\pi} e^{s(\theta)t} \tilde{f}(s(\theta)) s'(\theta)\, d\theta.$$

This integral is then approximated numerically using the trapezoidal rule, which is particularly effective for periodic integrands. The resulting discrete approximation takes the form

$$f(t) \approx \sum_{k=0}^{N-1} \omega_k\, e^{s_k t} \tilde{f}(s_k),$$

where the quadrature nodes $s_k = s(\theta_k)$ lie on the Talbot contour, with

$$\theta_k = -\pi + \frac{2\pi k}{N}, \qquad k = 0, \ldots, N-1,$$

and the weights ω_k incorporate both the trapezoidal coefficients and the derivative of the contour parameterization.

One of the key advantages of the Talbot method is its high accuracy with relatively few evaluation points. In practice, accurate results are typically obtained with modest values of N, making the method computationally efficient. Furthermore, the Talbot method is particularly well suited for functions arising in fractional calculus, where branch points and fractional powers frequently appear in the Laplace domain.

Another important feature of the Talbot method is its robustness with respect to different time scales. The contour scaling is adapted according to the value of t, ensuring stable computations for both short- and long-time responses. This property is especially relevant in viscoelastic modelling, where the material response may span several orders of magnitude in time.

In this work, the Talbot method [13] is therefore employed to compute the inverse Laplace transform of the relaxation modulus $G(t)$. These numerical inversions are performed repeatedly during the optimisation procedure, making computational efficiency essential.

3.2 Optimisation Problems

The parameters α, β, λ, $\mathbb{V}$, and g that define the proposed viscoelastic model are not known a priori and must therefore be estimated from experimental observations. In this work, these parameters are determined by solving optimisation problems that minimise the discrepancy between experimental data and the corresponding model predictions.

Two distinct experimental datasets are considered: (i) storage and loss moduli obtained in the frequency domain, and (ii) relaxation modulus obtained in the time domain. A two-step optimisation strategy is adopted here.

In the first step, the parameters are estimated using the storage and loss moduli data (denoted as *Problem 1*). This will result in the initial set of parameters (α_0, β_0, λ_0, $\mathbb{V}_0$, and g_0) to be used in the second step, where storage, loss and relaxation modulus data are used to obtain the optimal set of parameters (denoted as *Problem 2*).

– *(Problem 1):* The estimation of parameters α, β, λ, $\mathbb{V}$, and g is formulated as the following nonlinear least-squares optimisation problem:

$$\min_{\lambda,g,\alpha,\beta,\mathbb{V}\in\mathbb{R}} E_{G',G''}(\lambda,g,\alpha,\beta,\mathbb{V}) = \sum_{l=1}^{N_{G'G''}} \left(\left[\frac{\hat{G}'(\omega)_l}{G'_l} - 1 \right]^2 + \left[\frac{\hat{G}''(\omega)_l}{G''_l} - 1 \right]^2 \right)$$
$$\text{subject to} : \lambda > 0, g > 0, \mathbb{V} > 0,$$
$$\alpha, \beta \in (0,1).$$

Here, $N_{G'G''}$ denotes the total number of experimental observations available for the storage modulus G'_l and the loss modulus G''_l. The quantities $\hat{G}'(\omega)_l$ and $\hat{G}''(\omega)_l$ correspond to the values predicted by the model, evaluated at the angular frequency ω_l.

To solve *Problem 1*, the optimisation procedure is performed using the IPOPT solver, available through the NEOS server (http://neos.mcs.anl.gov/neos/solvers). The IPOPT (Interior Point OPTimizer) is designed for large-scale nonlinear programming problems and is based on an interior-point framework combined with a line-search filter strategy. This approach incorporates barrier methods together with filter-based acceptance criteria, allowing efficient handling of nonlinear constraints and improving convergence properties. Further details regarding the algorithmic structure and theoretical foundations can be found in [14,16–19].

For comparison purposes, the Generalized Reduced Gradient (GRG) nonlinear solver was also considered. This method optimises continuous parameters by iteratively computing descent directions that reduce the objective function. Consequently, the algorithm adjusts the model parameters so that the resulting smooth function provides a close agreement with the experimental data.

– *(Problem 2):* After obtaining the initial estimates α_0, β_0, λ_0, $\mathbb{V}_0$, and g_0, we proceeded by solving the following optimisation problem:

$$\min_{\lambda,g,\alpha,\beta,\mathbb{V}\in\mathbb{R}} E_{G',G'',G}(\lambda,g,\alpha,\beta,\mathbb{V}) = \sum_{l=1}^{N_{G'G''}} \left(\left[\frac{\hat{G}'(\omega)_l}{G'_l} - 1 \right]^2 + \left[\frac{\hat{G}''(\omega)_l}{G''_l} - 1 \right]^2 \right)$$
$$+ \sum_{l=1}^{N_G} \left[\frac{\hat{G}(t)_l}{G_l} - 1 \right]^2$$
$$\text{subject to} : \lambda > 0, g > 0, \mathbb{V} > 0,$$
$$\alpha, \beta \in (0,1).$$

In this formulation, N_G denotes the number of experimental data points for the relaxation modulus G_l, while $\hat{G}(t)_l$ represents the model prediction evaluated at time t_l.

It is important to note that the evaluation of $\hat{G}(t)$ requires the computation of an inverse Laplace transform. Therefore, we adopt a sequential solution strategy. In this approach, the optimisation algorithm operates only in the parameter space, while the inverse Laplace transform is computed separately at each iteration.

More specifically, for each candidate set of parameters, the inverse Laplace transform is computed using the Talbot method, followed by the evaluation of the objective function. This sequential procedure results in a bound-constrained optimisation problem that can be efficiently handled using standard nonlinear optimisation techniques.

The overall workflow of the sequential approach adopted for solving *Problem 2* is illustrated in Fig. 2.

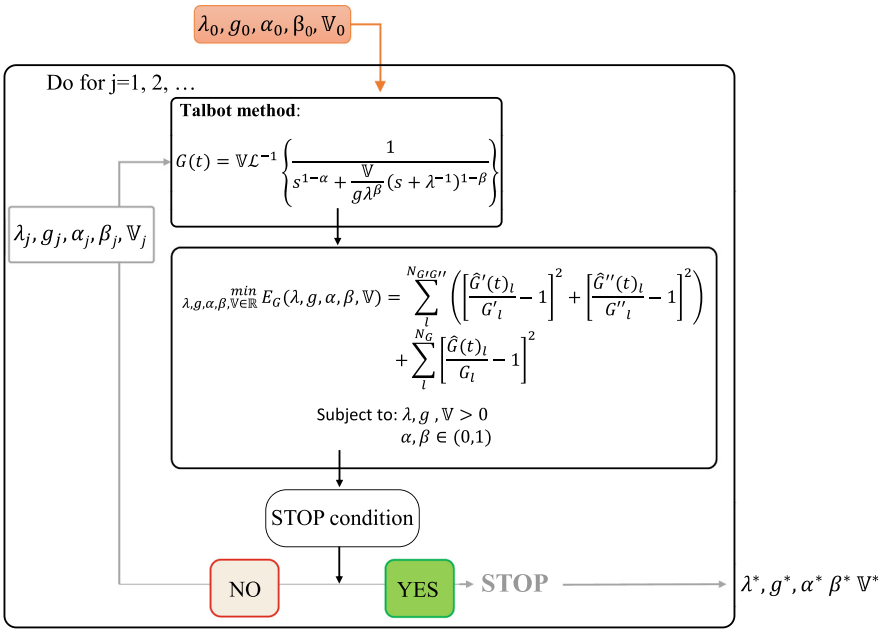


Fig. 2. Sequential algorithm used to solve *Problem 2*.

4 Fit to Data

We now present a fit to experimental data reported in [9] for the storage and loss modulus of a low-density polyethylene. The data were collected at different temperatures and shifted using the Williams–Landel–Ferry (WLF) equation

[11] to obtain a master curve. In addition, we consider experimental measurements of the relaxation modulus as a function of time for the same low-density polyethylene at 150°C.

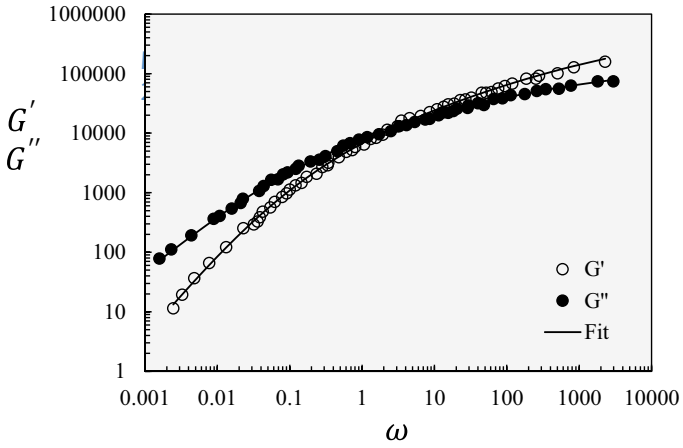


Fig. 3. Fit to experimental data reported in [9] for the storage and loss moduli (Pa) of a low-density polyethylene as functions of the angular frequency ω (rad/s).

The optimisation problem *Problem 1* was implemented using AMPL without presolve.

The optimisation problem was then solved using the IPOPT solver with default parameter settings. The solver was initialised with a set of starting values for the model parameters. After solving *Problem 1*, the following parameter estimates were obtained and subsequently used as initial guesses for *Problem 2*: $\alpha_0 = 0.51$, $\beta_0 = 0.15$, $\lambda_0 = 0.6$, $\mathbb{V}_0 = 14000$, and $g_0 = 80000$, yielding an objective function value of $E^*_{G',G''} = 0.5084$.

This optimal solution was achieved after 16 objective function evaluations, together with 16 evaluations of the objective gradient, 16 evaluations of the inequality constraints, 16 evaluations of the constraint Jacobian, and 16 evaluations of the Lagrangian Hessian. The total computational time reported by IPOPT (excluding function evaluations) was 0.05 CPU seconds, while the time spent on nonlinear programming function evaluations was 0.07 CPU seconds. All computations were carried out using the free NEOS cloud computing service.

To determine the final parameter values, *Problem 2* was solved using the sequential procedure illustrated in Fig. 2. This approach was implemented in Matlab®, where the inverse Laplace transform required for the evaluation of the relaxation modulus was computed numerically (the number of terms to sum for each t was 64 [12]) for each iteration.

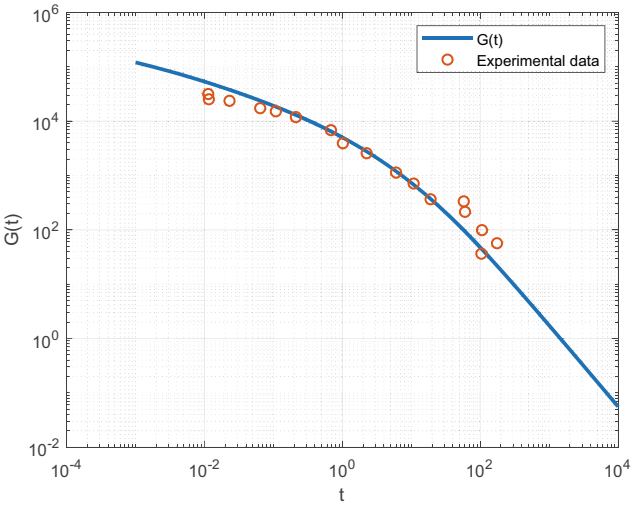


Fig. 4. Fit to experimental data reported in [9] for the relaxation modulus $G(t)$ (Pa) of a low-density polyethylene as a function of time t (s).

The optimisation was performed using the *fmincon* function with the interior-point algorithm. A stopping criterion based on a tolerance of 1×10^{-10} was adopted, while all remaining solver parameters were kept at their default values.

The optimisation procedure converged after 10 iterations, leading to the following optimal parameter values: $\alpha_0 = 0.4984$, $\beta_0 = 0.1513$, $\lambda_0 = 0.5678$, $\mathbb{V}_0 = 13084.8022$, and $g_0 = 90309.8388$.

The resulting global error was $E_{G',G'',G} = 0.6126$. This relatively small error indicates that the proposed model provides an accurate representation of the experimental data. In particular, a very good agreement was obtained for both the storage and loss moduli (Fig. 3) and the relaxation modulus (Fig. 4).

5 Conclusions

In this work, a fractional viscoelastic model composed of a springpot connected in series with a fractional dashpot was proposed, together with a systematic strategy for parameter identification based on experimental rheological data. Fractional models of this type offer a flexible and compact framework for representing complex viscoelastic behaviour using a limited number of physically meaningful parameters. In particular, the proposed configuration is capable of capturing both short- and long-time memory effects, which are frequently observed in polymeric materials and are often difficult to reproduce using classical integer-order constitutive models.

The model parameters were identified using experimental data obtained in both the frequency and time domains, namely the storage modulus, the loss modulus, and the relaxation modulus. These complementary datasets provide a

good characterisation of the viscoelastic response and enable a more reliable estimation of the model parameters. To enhance both robustness and computational efficiency, a two-step optimisation strategy was adopted.

In the first stage, parameter estimation was performed by fitting the storage and loss modulus data. The corresponding optimisation problem was formulated as a nonlinear least-squares problem, in which the objective function measures the discrepancy between experimental observations and model predictions over a range of angular frequencies. This problem was solved using the IPOPT solver, which is well suited for nonlinear optimisation problems with bound constraints. The resulting parameter estimates were stable and physically meaningful, and were subsequently used as initial guesses for the second stage.

In the second stage, the optimisation procedure incorporated both frequency-domain data and time-domain data through the relaxation modulus. Since the relaxation modulus is obtained by numerically inverting Laplace-domain expressions, the resulting optimisation problem becomes computationally more demanding. To address this difficulty, a sequential approach was adopted, whereby the numerical inversion using the Talbot method and the optimisation procedure were performed iteratively until convergence was achieved.

The results demonstrate that the proposed fractional model provides an excellent agreement with the experimental data in both the frequency and time domains. In particular, the model accurately reproduces the storage and loss moduli over the entire frequency range considered, as well as the relaxation modulus across the time interval analysed. These findings confirm the capability of the springpot–fractional dashpot configuration to capture complex viscoelastic behaviour, including power-law responses and long-memory effects characteristic of fractional-order models.

Moreover, the proposed optimisation strategy proved to be robust and computationally efficient, allowing reliable parameter estimation even when numerical inversion of Laplace-domain expressions is required.

A limitation of the present work is that the parameter identification and validation procedures were carried out using experimental data corresponding only to a single material, namely low-density polyethylene. Although the obtained results demonstrate the suitability of the proposed fractional viscoelastic model for this polymer, the general applicability of the model to other classes of viscoelastic materials remains to be investigated. In particular, materials exhibiting different rheological characteristics, such as highly filled polymers, biological tissues, elastomers, or composite materials, may require different fractional structures or additional constitutive elements to accurately capture their behaviour. Furthermore, the identification procedure considered only three types of experimental data: the storage modulus, the loss modulus, and the relaxation modulus. While these quantities provide important information regarding the linear viscoelastic response, other experimental measurements, such as creep compliance, stress–strain, or nonlinear rheological tests, were not included and could provide additional validation of the model and its predictive capabilities.

Future work may therefore focus on extending the proposed methodology to a broader range of viscoelastic materials and experimental conditions. In par-

ticular, it would be of interest to investigate the performance of the model for materials exhibiting nonlinear or temperature-dependent behaviour, as well as to study the robustness of the parameter identification procedure under noisy or incomplete experimental data. Additional developments may also include the incorporation of other rheological tests, such as creep and dynamic loading experiments, into the optimisation framework in order to improve parameter identifiability and model reliability. From a computational perspective, alternative numerical inversion techniques and more advanced optimisation strategies could be explored to further reduce computational cost and improve convergence properties. Finally, the extension of the proposed fractional framework to multidimensional constitutive models and its implementation in numerical simulation methods, would constitute an important step toward practical engineering applications.

Acknowledgments. We thank the support of the High-Performance Computing Center at the University of Évora funded by FCT I.P. under the project "OptXAI: Constrained optimisation in NNs for Explainable, Ethical and Greener AI", reference 2024.00191.CPCA.A1, platform Vision. This work was financially supported by national funds through the FCT/MCTES (PIDDAC) under the project 2022.06672.PTDC—iMAD (Improving the Modelling of Anomalous Diffusion and Viscoelasticity: solutions to industrial problems), DOI 10.54499/2022.06672.PTDC (https://doi.org/ 10.54499/ 2022.06672.PTDC); and by the project UID/00013/2025 (https://doi.org/10.54499/UID/00013/2025) (CMAT). It was also financially supported by Fundação "la Caixa"|BPI and FCT through project PL24-00057: "Inteligência Artificial na Otimização da Rega para Olivais Resilientes às Alterações Climáticas". C. Coelho would like to thank the KIBIDZ project funded by dtec.bw-Digitalization and Technology Research Center of the Bundeswehr; dtec.bw is funded by the European Union-NextGenerationEU.

M.L. Morgado acknowledges funding by FCT - Fundação para a Ciência e a Tecnologia, through the project https://doi.org/10.54499/UID/04621/2025.

L.L. Ferrás acknowledges FCT - Fundação para a Ciência e a Tecnologia, for financial support through FCT/MECI: CEFT, UID/00532/2025 (https://doi.org/10.54499/UID/ 00532/2025) and UID/PRR/ 00532/2025 (https://doi.org/10.54499/UID/PRR/ 00532/2025) and ALiCE, LA/P/0045/2020 (https:// doi.org/10.54499/LA/P /0045 /2020).

M. Rebelo acknowledges FCT – Fundação para a Ciência e a Tecnologia, I.P., under the scope of the projects UID/00297/2025 (https://doi.org/10.54499/UID/00297/ 2025) and UID/PRR/00297/2025 (https://doi.org/10.54499/UID/PRR/00297/2025) (Center for Mathematics and Applications – NOVA Math).

References

1. Bird, R.B., Armstrong, R.C., Hassager, O.: Dynamics of Polymeric Liquids. Vol. 1: Fluid Mechanics. A Wiley-Interscience Publication, John Wiley & Sons (1987)
2. Diethelm, K.: The Analysis of Fractional Differential Equations. Lecture Notes in Mathematics, vol. 2004. Springer, Berlin, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14574-2

3. Ferrás, L.L., et al.: Theoretical and numerical analysis of unsteady fractional viscoelastic flows in simple geometries. Comput. Fluids **174**, 14–33 (2018)
4. Ferrás, L.L., Luísa, M., Rebelo, M., Mckinley, G.H., Afonso, A.: A generalized phan-thien – tanner model. J. Nonnewton. Fluid Mech. **269**, 88–99 (2019)
5. Ferrás, L.L., Morgado, M.L., Rebelo, M.: A numerical method for the transient couette flow of a distributed-order viscoelastic fluid. In: Gervasi, O., et al. (eds.) Computational Science and Its Applications – ICCSA 2021, Lecture Notes in Computer Science, vol. 12952. Springer, Cham (2021)
6. Ferrás, L.L., Morgado, M.L., Rebelo, M.: A generalised distributed-order maxwell model. Math. Methods Appl. Sci. **46**(1), 368–387 (2023)
7. Ferrás, L.L., Rebelo, M., Morgado, M.L.: The role of the weight function in the generalised distributed-order maxwell model: the case of a distributed-springpot and a dashpot. Appl. Math. Model. **122**, 844–860 (2023)
8. Henriques, R.O., Ferrás, L.L., Rebelo, M., Morgado, L.M.: Mathematical analysis of a novel linear fractional viscoelastic model exhibiting fluid-like behaviour (2026), manuscript submitted for publication
9. Laun, H.M.: Description of the non-linear shear behaviour of a low density polyethylene melt by means of an experimentally determined strain dependent memory function. Rheol. Acta **17**(1), 1–15 (1978)
10. Leiva, R.T., Ferrás, L.L., Castelo, A., et al.: A generalisation of the integral maxwell model: the GK-BKZ model–frame invariance and analytical solutions. Meccanica **59**, 363–384 (2024)
11. Macosko, C.W.: Rheology: Principles, Measurements, and Applications. Wiley-VCH, New York (1994)
12. McClure, T.: Numerical inverse laplace transform. https://www.mathworks.com/matlabcentral/fileexchange/39035-numerical-inverse-laplace-transform (2026). Retrieved: 19 Mar 2026
13. Murli, A., Rizzardi, M.: Algorithm 682: Talbot's method for the laplace inversion problem. ACM Trans. Math. Softw. **16**(2), 158–168 (1990)
14. Nocedal, J., Wächter, A., Waltz, R.A.: Adaptive barrier strategies for nonlinear interior methods. SIAM J. Optim. **19**(4), 1674–1693 (2009)
15. Schiessel, H., Metzler, R., Blumen, A., Nonnenmacher, T.F.: Generalized viscoelastic models: their fractional equations with solutions. J. Phys. A: Math. Gen. **28**, 6567–6584 (1995)
16. Wächter, A.: An Interior Point Algorithm for Large-Scale Nonlinear Optimization with Applications in Process Engineering. Ph.D. thesis, Carnegie Mellon University, Pittsburgh, PA, USA (2002)
17. Wächter, A., Biegler, L.T.: Line search filter methods for nonlinear programming: local convergence. SIAM J. Optim. **16**(1), 32–48 (2005)
18. Wächter, A., Biegler, L.T.: Line search filter methods for nonlinear programming: motivation and global convergence. SIAM J. Optim. **16**(1), 1–31 (2005)
19. Wächter, A., Biegler, L.T.: On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming. Math. Program. **106**(1), 25–57 (2006)
20. Yao, D.: A fractional dashpot for nonlinear viscoelastic fluids. J. Rheol. **62**(2), 619–629 (2018)

Chaotic System Fingerprinting for Cryptographic Keystream Attribution: A Deep Learning Approach

Alexandru Dinu, Sara-Gabriela Modan, and Mădălin Frunzete(✉)

National University of Science and Technology POLITEHNICA Bucharest, Faculty of Electronics, Telecommunications and Information Technology, Bucharest, Romania
{alexandru.dinu2408,madalin.frunzete}@upb.ro,
sara.gabriela.modan@stud.etti.upb.ro

Abstract. Chaotic dynamical systems are widely used as entropy sources for key generation used for cryptographic purposes. This introduces a vulnerability: if an adversary were able to identify which chaotic system produced a given keystream, they would gain an important advantage for cryptanalysis. In this paper, we propose a deep learning framework for chaotic system fingerprinting, formulated in terms of a time-series classification problem over short keystream sequences. The model is trained and evaluated on sequences generated by four representative chaotic systems (Lorenz, Rössler, Sprott, Hénon 3D), using strict train/test separation. A regularized Residual Network (ResNet) model achieves over 92% accuracy on 4-level quantized sequences and over 86% on binary sequences, performing significantly better than frequency-based and classical Machine Learning (ML) baselines. An important observation is that these results hold even though the sequences pass standard statistical randomness tests which could make one think the classification results should be less positive. This demonstrates that high entropy does not prevent system identification. As a consequence, key generators based on chaotic systems remain vulnerable to system-level identification and require cryptographic post-processing regardless of their statistical properties.

Keywords: Chaotic systems · Deep learning · Cryptographic fingerprinting · ResNet · Security analysis

1 Introduction

Chaotic systems like Lorenz, Rössler, Sprott and Hénon are very attractive for cryptography because they are able to produce pseudo-random sequences from simple deterministic rules [1–4]. Their sensitivity to initial conditions creates large key spaces, and their outputs typically have no issues in passing standard randomness tests. This has led to widespread use in stream ciphers, image encryption, and secure communication, often under the (implicit) assumption

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 298–312, 2027.
https://doi.org/10.1007/978-3-032-30530-5_19

that statistical randomness is sufficient for security. However, the evaluation of such systems has in most of the cases relied on statistical criteria, such as entropy and standard randomness test suites, rather than on resistance to inference or attribution attacks. This leaves open the question of whether seemingly random outputs may still encode structural information about their generating dynamics.

A key assumption is behind the usage of chaos in cryptography: observing a keystream reveals nothing about which chaotic system generated it. If this assumption is not valid—i.e., if Lorenz leaves significantly different fingerprints/traces compared to Rössler or Sprott—an adversary who correctly identifies the source gains a concrete advantage. When an attacker knows the system used, this can enable targeted attacks: exploiting system-specific statistical weaknesses [6], making use of characteristic temporal correlations for parameter estimation, or applying phase space reconstruction specific to the identified attractor.

Even though there are rich decades of chaos-based cryptography research [3, 4], the question regarding the classification and correct identification of chaotic systems has not been systematically studied. Prior work focused on optimizing individual systems for statistical randomness and synchronization purposes. However, if different attractors produce outputs that a classifier can reliably differentiate did not receive the attention it deserves until now.

In the present paper, we address this topic in a direct manner. The question we started from is: "Given a short quantized sequence from an unknown chaotic source, can a deep learning model identify which attractor produced it?" We show that it can. Furthermore, this holds even for sequences that pass standard randomness tests and exhibit almost maximum informational entropy. A further finding concerns quantization: the same attractor can behave very differently depending on how its output is discretized (2 or 4 levels), yet remains equally identifiable under both use cases. These results suggest that entropy alone is not sufficient to assess the security of chaos-based key generators, and that cryptographic post-processing is necessary even for systems that appear statistically strong.

The remainder of this paper is organized as follows:

- Sect. 2 reviews related work.
- Sect. 3 describes the dataset, models, and experimental setup.
- Sect. 4 presents the results.
- Sect. 5 discusses security implications
- Sect. 6 presents the conclusions of this work.

2 Background and Related Work

2.1 Chaotic Dynamical Systems

Chaotic systems are deterministic nonlinear systems exhibiting extreme sensitivity to initial conditions [1]. A system is chaotic if it possesses at least one positive

Lyapunov exponent, indicating exponential divergence of trajectories which are close to each other. The four attractors studied here are defined as follows.

The Lorenz system (3D continuous, $\sigma = 10$, $\rho = 28$, $\beta = 8/3$):

$$\frac{dx}{dt} = \sigma(y - x), \quad \frac{dy}{dt} = x(\rho - z) - y, \quad \frac{dz}{dt} = xy - \beta z \tag{1}$$

exhibits a butterfly-shaped strange attractor with two lobes.

The Rössler system (3D continuous, $a = 0.2$, $b = 0.2$, $c = 5.7$):

$$\frac{dx}{dt} = -y - z, \quad \frac{dy}{dt} = x + ay, \quad \frac{dz}{dt} = b + z(x - c) \tag{2}$$

features a single-lobed attractor with folded ribbon topology.

The Sprott system (3D continuous, $a = 2.07$, $b = 1.79$):

$$\frac{dx}{dt} = y + axy + xz, \quad \frac{dy}{dt} = 1 - bx^2 + yz, \quad \frac{dz}{dt} = x - x^2 - y^2 \tag{3}$$

is time-reversible, with a state-dependent divergence that leads to both locally conservative and dissipative behavior across different regions of the phase space.

The Hénon 3D map (discrete-time, $a = 1.76$, $b = 0.1$):

$$x_{n+1} = a - y_n^2 - bz_n, \quad y_{n+1} = x_n, \quad z_{n+1} = y_n \tag{4}$$

is a three-dimensional extension of the classical 2D Hénon map, introducing delayed feedback through $z_n = y_{n-1}$. Positive Lyapunov exponents were verified numerically, confirming chaotic behavior suitable for cryptographic applications [1,7–10].

Figure 1 shows the phase portraits of all four systems. Each attractor has a structure which is distinct from a geometric point of view, with trajectories that start close to each other and gradually become distinct, but remain confined to the same attractor.

2.2 Chaos-Based Cryptography

Chaos-based cryptographic systems generate keystreams by quantizing trajectories from chaotic attractors which start from an initial condition and a clear set of system parameters [2,3]. A simple encryption scheme works by pairing each plaintext with a key value generated from the chaotic system:

$$C_i = M_i \oplus K_i, \quad \text{where } K_i = \text{quantize}(x_i).$$

In this approach, the continuous output of the chaotic system is first converted into discrete values, which are then used as the keystream.

Security analysis has traditionally focused on several security properties like: large key space, statistical randomness via NIST tests [5], and unpredictability from positive Lyapunov exponents [6]. Whether an attacker can establish which attractor produced a given keystream—the inter-system perspective—has not been sufficiently explored until now. This gap is important: successful fingerprinting enables targeted exploitation of system-specific weaknesses rather than generic cryptanalysis.

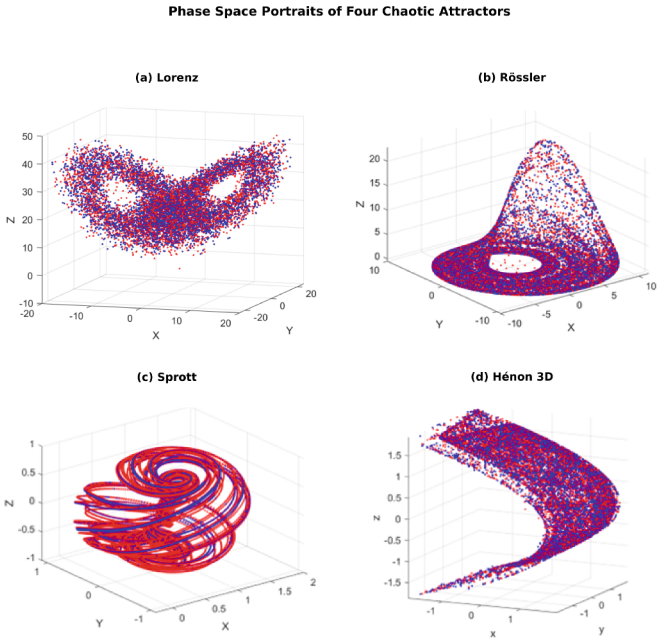


Fig. 1. Phase space portraits of the four chaotic attractors studied: (a) Lorenz, (b) Rössler, (c) Sprott, and (d) Hénon 3D.

2.3 Deep Learning for Time-Series and Cryptographic Attacks

Residual networks (ResNets) address vanishing gradients through skip connections, enabling deep architectures for temporal feature extraction via 1D convolutions [11]. Chaotic sequences present specific challenges [12]: short length and limited training data make overfitting severe, requiring aggressive regularization and extended training patience.

In the cryptographic attack literature, Machine Learning (ML) has been applied to problems such as modeling physical unclonable functions (PUFs), recovering the internal state of pseudo-random number generators (PRNGs), and identifying algorithms through side-channel leakage [13,14]. PUFs refer to hardware-based fingerprints that are difficult to replicate, PRNGs are deterministic generators designed to produce sequences that appear random, and side-channel attacks exploit unintended information leakage (such as timing or power consumption) from implementations.

Our work differs by targeting the mathematical structure of the underlying primitive—the attractor type—from outputs alone, without relying on physical side channels.

3 Methodology

3.1 Dataset Generation and Preprocessing

The data used was generated as follows:

1. Continuous systems were simulated using MATLAB's `ode45` solver, which provides a high-accuracy numerical approximation of the underlying differential equations. The step size is adaptive, ensuring numerical stability and reducing discretization artifacts.
2. The discrete Hénon 3D map trajectories were obtained through direct iteration.
3. Neural network models were implemented in Python using TensorFlow/Keras. No pre-trained weights were used; all models were trained from scratch on the generated dataset.

For each system, we generated $N = 1{,}000$ independent trajectories from random initial conditions $\mathbf{x}_0 \sim \mathcal{N}(0, I_3)$.

Statistical validity was ensured through two-stage validation: transient detection via histogram stabilization to identify attractor convergence time t_{trans}, and temporal independence testing via the U-V scatter chi-square analysis proposed by Badea & Vlad and used by the authors in previous research [8] to determine the minimum lag $\Delta t_{\min}$ satisfying

$$z = \sum_{k=1}^{M} \frac{(M_k - N/M)^2}{N/M} < \chi^2_{\alpha=0.05,\, M-1} \tag{5}$$

where samples were partitioned into $M = 48$ sectors (6 rings $\times$ 8 wedges). Table 1 summarizes the validated parameters. Hénon requires 1,300 s to reach its attractor (65$\times$ Lorenz), while Sprott demands 350 s spacing (22$\times$ Lorenz).

Table 1. Transient periods and minimum independence lags per system.

System	t_{trans} (s)	$\Delta t_{\min}$ (s)
Lorenz	20	16
Rössler	100	40
Sprott	100	350
Hénon 3D	1,300	200

After discarding transients, we extracted 150-timestep sequences by sampling the x-coordinate at intervals $\Delta t_{\min}$, ensuring temporal independence both within and across sequences.

For 4-level quantization, per-sequence uniform binning into $\{0, 1, 2, 3\}$ was applied:

$$s_i^{\text{4-level}} = \left\lfloor 4 \cdot \frac{s_i - \min(s_j)}{\max(s_j) - \min(s_j)} \right\rfloor \tag{6}$$

For binary quantization, per-sequence median thresholding produces balanced 0/1 distributions:

$$s_i^{\text{bin}} = \begin{cases} 0 & \text{if } s_i \leq \text{median}(\{s_j\}) \\ 1 & \text{if } s_i > \text{median}(\{s_j\}) \end{cases} \tag{7}$$

Per-sequence normalization ensures that temporal patterns, not amplitude scales, drive classification.

The 4-level dataset contains 726 sequences (580 train, 146 test). The binary dataset (344 sequences: 275 train, 69 test) is smaller because we retained only sequences satisfying $|N_0 - N_1| \leq 10$, as a simple balancing condition that ensures the number of zeros and ones remains approximately equal, improving cryptographic quality. Table 2 gives the class breakdown. While the dataset is not perfectly balanced across classes, the class distributions were kept consistent across training and testing splits. This reflects a realistic setting and allows the model to learn discriminative features without enforcing artificial uniformity. The source code is available upon request from the authors.

Table 2. Dataset composition for chaotic system classification.

System	4-Level		Binary	
	Train	Test	Train	Test
Lorenz	154	38	55	14
Rössler	151	38	101	25
Sprott	122	31	65	16
Hénon 3D	153	39	54	14
Total	580	146	275	69

3.2 Model Architectures

The four systems were selected to cover distinct attractor geometries and dynamical classes: Lorenz and Rössler are classical continuous-time systems with established cryptographic relevance [1,7], Sprott is a time-reversible continuous-time variant, and Hénon 3D is the only discrete-time map, included to probe whether the classifier exploits the continuous/discrete distinction. The architectures follow the benchmark protocol of Fawaz et al. [12] for time-series classification, ranging from a simple dense baseline to deep residual networks. We evaluate four architectures of increasing complexity on 150-timestep sequences.

Dense Simple (27.8K parameters) is a fully-connected network with two hidden layers (128 and 64 units, ReLU, dropout 0.3). It tests whether aggregate statistics alone are sufficient, without exploiting temporal structure.

Convolutional Neural NetworkS (CNN) Light (100.7K parameters) uses three Conv1D layers (64, 128, 128 filters; kernels 7, 5, 3), max pooling, batch normalization, and dropout 0.25.

ResNet Baseline (523.6K parameters) adds residual connections: three blocks (64, 128, 128 filters), each with three convolutional layers (kernels 8, 5, 3), batch normalization, ReLU, and skip connections via 1×1 convolution. It uses moderate regularization (L2=0.0005, dropout 0.2–0.3) and early stopping with learning rate reduction.

ResNet Robust uses three residual blocks (filters: 64, 128, 128), each containing three 1D convolutions with kernel sizes 8, 5, and 3 respectively, followed by batch normalization. Each block includes a shortcut connection (1×1 convolution + batch normalization). Figure 2 illustrates the architecture: input sequences (150×1) pass through an initial convolutional block and three residual blocks, followed by global average pooling and a dense classification head with dropout.

Both ResNet variants use Adam (lr=0.001) with stratified 5-fold cross-validation.

Training Strategy

All models use stratified 5-fold cross-validation, trained from scratch on four folds and evaluated on the held-out fold. We report mean accuracy and standard deviation across five runs. Training uses Adam with initial learning rate 0.001 and sparse categorical cross-entropy loss, with EarlyStopping (patience: 10 epochs for Dense, 20 for CNN and ResNet Baseline, 30 for ResNet Robust) and ReduceLROnPlateau (factor 0.5–0.7, patience 7–10 epochs). Batch sizes are 16 for Dense and ResNet Robust, 32 for CNN and ResNet Baseline.

4 Experimental Results

4.1 Overall Classification Performance

Table 3 shows cross-validation results on the 4-level quantized dataset. ResNet Baseline reaches the highest mean accuracy, while ResNet Robust trades a small amount of accuracy for substantially lower variance across folds—standard deviation drops by 28%. Both Dense Simple and CNN Light fall well short of the ResNet variants.

Table 3. Model performance on 4-level quantized dataset (726 sequences, 5-fold CV).

Model	Accuracy (%)	Macro F1	Params	Epochs
ResNet Baseline	**93.25 ± 1.98**	**0.933 ± 0.020**	523.6K	55.8 ± 9.7
ResNet Robust	**92.83 ± 1.43**	**0.928 ± 0.015**	523.6K	64.2 ± 12.4
CNN Light	73.83 ± 3.37	0.726 ± 0.041	100.7K	21.0 ± 0.0
Dense Simple	71.77 ± 2.01	0.719 ± 0.020	27.8K	20.6 ± 1.4

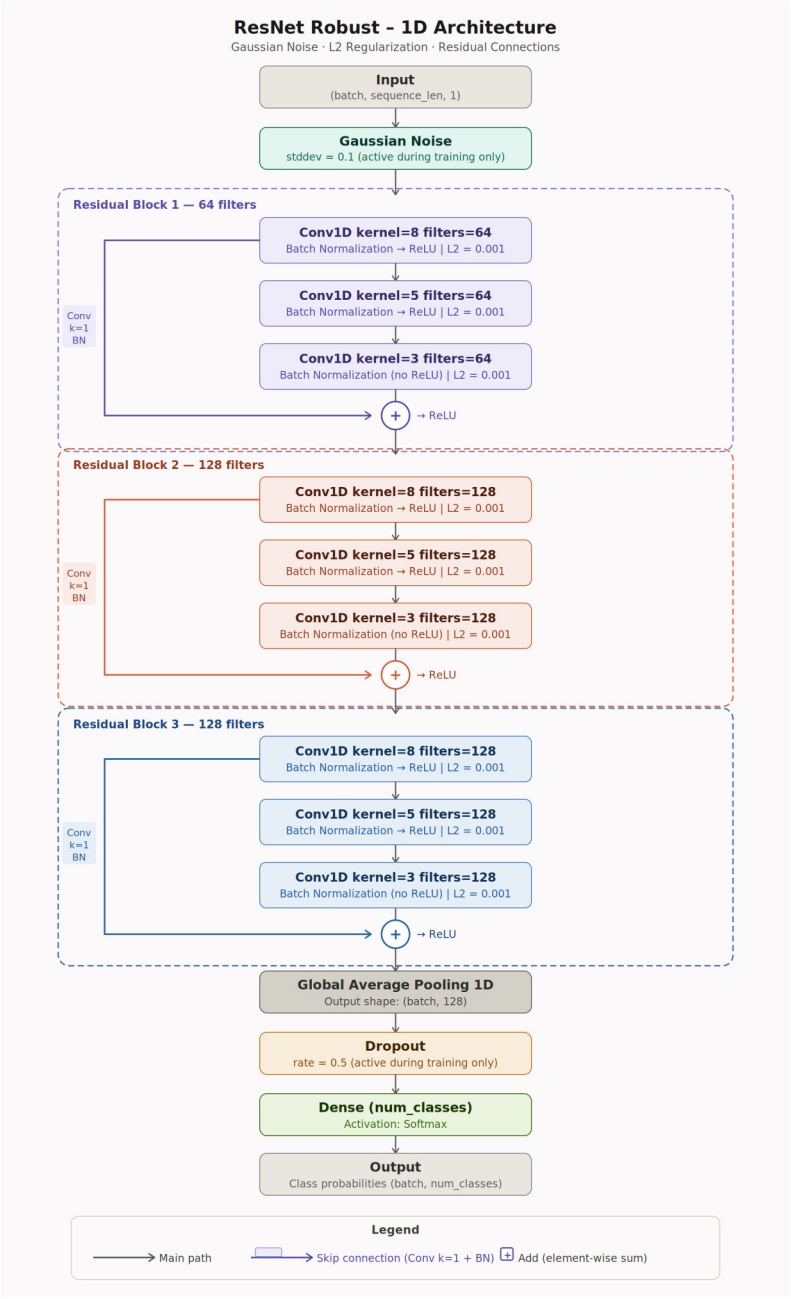


Fig. 2. ResNet Robust architecture for chaotic time-series classification (523,588 parameters).

On binary data the picture changes considerably. Table 4 shows that ResNet Baseline, despite being the best performer on 4-level data, collapses to near-random accuracy. With 523K parameters and only 275 training samples, it memorizes training noise within a few epochs. ResNet Robust, using the identical architecture but with stronger regularization, maintains high accuracy. The difference is entirely due to training strategy: Gaussian noise at input, stronger L2 weight decay, higher dropout, and extended patience allow the network to converge slowly and learn generalizable patterns rather than fit noise.

Table 4. Model performance on binarized dataset (344 sequences, 5-fold CV).

Model	Accuracy (%)	Macro F1	Params	Epochs
ResNet Robust	**86.34 ± 2.68**	**0.859 ± 0.026**	523.6K	92.0 ± 6.8
Dense Simple	59.59 ± 5.39	0.563 ± 0.061	27.8K	19.2 ± 2.0
CNN Light	51.46 ± 9.30	0.335 ± 0.064	100.7K	23.0 ± 0.0
ResNet Baseline	48.23 ± 8.83	0.357 ± 0.089	523.6K	24.2 ± 5.0

Table 5 summarizes how each model behaves when the representation is reduced from 4-level to binary quantization. As expected, this reduction leads to a loss of information and less training data. However, the ResNet Robust model maintains most of its performance, with only a 7% decrease. At the same time, the difference of 38.53% points between the two identically structured ResNet variants shows that regularization alone plays a major role, since all other factors are kept the same.

Table 5. Accuracy degradation from 4-level to binary quantization.

Model	4-Level (%)	Binary (%)	Δ (pp)	Δ (%)
ResNet Robust	**92.83 ± 1.43**	86.34 ± 2.68	−6.49	−7.0
Dense Simple	71.77 ± 2.01	59.59 ± 5.39	−12.18	−17.0
CNN Light	73.83 ± 3.37	51.46 ± 9.30	−22.37	−30.3
ResNet Baseline	**93.25 ± 1.98**	48.23 ± 8.83	−45.02	−48.3

4.2 Per-Class Analysis

Table 6 shows the F1-scores for each attractor. On the 4-level dataset, all four systems are classified with high reliability. The Hénon 3D system is the easiest to identify: both ResNet variants achieve 100% recall, meaning they do not miss any Hénon sequences. It should be noted that Hénon 3D is the only discrete-time system among the four; the remaining three are continuous-time ODEs.

The classifier may therefore partly exploit the continuous/discrete distinction rather than purely attractor identity. This does not invalidate the fingerprinting result, since both are real structural properties, but future work should evaluate classifiers restricted to sets of exclusively continuous-time or exclusively discrete-time attractors to isolate attractor-specific signatures.

Table 6. Per-class F1-scores (mean ± std, %) for ResNet architectures.

System	4-Level Quantized		Binary	
	Baseline	Robust	Baseline	Robust
Lorenz	90.3 ± 2.4	91.5 ± 2.0	18.8 ± 23.0	89.1 ± 6.7
Rössler	88.9 ± 3.0	89.0 ± 2.3	45.9 ± 25.1	86.2 ± 3.1
Sprott	95.5 ± 2.2	93.7 ± 1.6	74.2 ± 12.3	94.9 ± 4.5
Hénon 3D	98.7 ± 0.8	97.0 ± 2.0	4.1 ± 8.2	73.5 ± 5.1

Binary quantization reveals a clear difference between the two architectures. The ResNet Baseline model performs very poorly on the Lorenz and Hénon systems, and its high variability across folds suggests that its behavior is unstable and depends strongly on how the data is split, rather than on meaningful learned patterns. In contrast, the ResNet Robust model maintains performance above 73% F1 for all four systems, with low variance (under 7%), indicating consistent and stable learning. This shows that regularization does more than just prevent the model from collapsing to a trivial solution—it helps the network learn all classes instead of focusing too early on the dominant patterns during training.

Figure 3 shows confusion matrices from the best-performing fold. On 4-level data, the strong diagonal confirms reliable learning across all four systems, with Hénon 3D perfectly recalled. On binary data, Lorenz and Sprott achieve perfect precision in this fold, while Hénon 3D drops to 61.5% recall—still more than double the 25% random baseline.

4.3 Entropy vs. Classifiability

Table 7 compares Shannon entropy against classification performance on 4-level quantized data. Entropy values are computed across all sequences for each system (Lorenz: 192, Rössler: 189, Sprott: 153, Hénon: 192) after binarization via the median thresholding procedure already described.

Lorenz achieves near-perfect entropy and passes the chi-square uniformity test, yet remains highly classifiable. Hénon 3D reaches the highest F1-score of any system despite also passing randomness tests, with 100% recall in multiple folds. Rössler fails chi-square tests and shows high temporal correlation (lag-1 autocorrelation: 0.262), making it both cryptographically weak and easily identifiable—a dual failure. Sprott has the lowest entropy with extreme autocorrelation (lag-1:

Table 7. Shannon entropy versus classification performance on 4-level quantized data (post-binarization).

System	Entropy (H)	F1-Score (%)	Gap to Random	Chi^2 p	Status
Lorenz	0.9999	91.5 ± 2.0	+66.5 pp	0.123	Vulnerable
Rössler	0.9344	89.0 ± 2.3	+64.0 pp	$< 10^{-300}$	Critical
Sprott	0.7377	93.7 ± 1.6	+68.7 pp	$< 10^{-300}$	Critical
Hénon 3D	0.9812	97.0 ± 2.0	+72.0 pp	$< 10^{-164}$	Highly Vulnerable

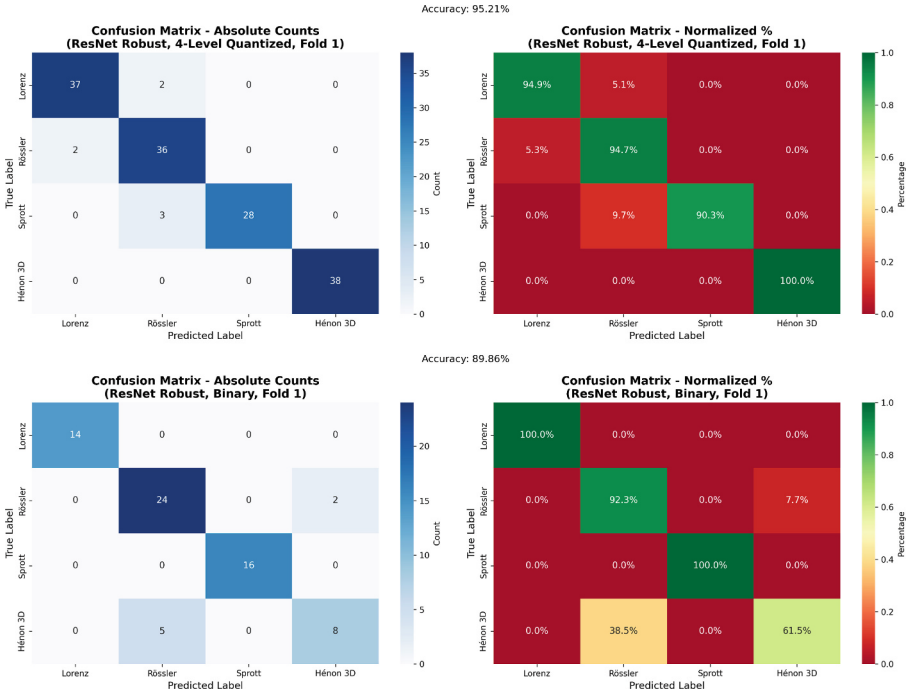


Fig. 3. Confusion matrices for ResNet Robust on the highest-performing cross-validation fold (shown for illustration); mean performance across folds is reported in Tables 3 and 4.

0.617) and a longest run of 162 consecutive identical bits out of 22,950 total, yet is classified just as reliably as Lorenz.

The reason standard randomness tests fail to capture this effect is that they focus on simple statistical properties of the bit sequences, such as frequency and uniformity. In contrast, deep learning models can detect more complex structures, including temporal dependencies and patterns that persist after quantization. As a result, even sequences that appear random under classical tests still contain identifiable signatures of their underlying dynamics.

Table 8 shows the same analysis on native binary data, where the impact of the quantization scheme becomes clear.

Table 8. Shannon entropy on native binary data (no median thresholding).

System	Entropy (H)	Δ vs Quant4	Chi^2 p	Runs p
Lorenz	0.9998	−0.0001	0.120	0.941
Rössler	0.9485	+0.0141	$< 10^{-292}$	$< 10^{-300}$
Sprott	0.9782	+0.2405	$< 10^{-81}$	$< 10^{-300}$
Hénon 3D	0.9897	+0.0085	$< 10^{-33}$	0.065

Sprott's entropy jumps from 0.7377 (post-binarized quant4) to 0.9782 (native binary)—a gain of 24% points. Median-threshold binarization of 4-level quantized Sprott sequences degrades entropy severely, while direct binary quantization preserves attractor structure. Despite this recovery, Sprott still fails statistical tests and ResNet Robust classifies both variants with comparable accuracy. Rössler recovers only modestly (+1.41 pp), continuing to fail uniformity tests with persistent temporal correlation. Lorenz and Hénon remain statistically robust across both schemes.

In practice, this means that evaluating entropy for a single quantization method is not enough. A chaos-based cryptosystem may appear secure under one discretization, but reveal significant weaknesses under another. At the same time, it can remain equally vulnerable to ML-based fingerprinting in both cases.

5 Security Implications

5.1 Threat Model

Consider an adversary who intercepts a keystream and aims to infer which chaotic system generated it. While the exact system is not known, the attacker may assume a set of commonly used candidates and train a classifier on representative data. Our results show that this inference is feasible with high accuracy, with individual folds exceeding 95%. Identifying the underlying system provides a concrete advantage, enabling targeted cryptanalysis based on system-specific statistical properties and dynamical behavior. It also facilitates parameter estimation and offers insight into the defender's cryptographic design choices.

The main way to address this issue is to apply cryptographic post-processing. Even if the generated sequences look highly random, they can still be identified if used directly. By hashing the output (for example, with SHA-256), the structural patterns specific to each chaotic system are removed, transforming the output into a sequence that is computationally indistinguishable from uniform random noise, and removing the structural signatures that the fingerprinting classifier exploits.

Designers must also validate randomness across multiple quantization schemes, since discretization choices are not cryptographically neutral—vulnerabilities can emerge from preprocessing decisions that appear innocuous. Testing a single pipeline is not sufficient. Additional defensive layers include multi-attractor hybrids [8] (e.g., $K_{\text{hybrid}} = K_{\text{Lorenz}} \oplus K_{\text{Hénon}}$), which make attribution harder even when individual components are fingerprintable, and one-time initial conditions to prevent adversaries from accumulating training data across sessions.

System-specific guidance: Hénon 3D should not be used as a sole entropy source. Lorenz can be deployed if paired with SHA-256 post-processing. Rössler and Sprott are not recommended without post-processing—both exhibit severe statistical failures in addition to high ML classifiability.

Discussion

The main observation which results from this study is that statistical randomness does not imply security against ML-based classification. A sequence can pass every standard randomness test and still be classified in a reliable manner based on its source attractor. The rationale is chaos's dual character: locally the output looks unpredictable, but globally the dynamics follow deterministic geometric patterns in phase space that deep networks can extract.

Training strategy is as important as the architecture. Two neural networks with identical structure—same parameters, same architecture—can behave very differently on scarce data depending solely on regularization strength. Aggressive regularization forces slower convergence and prevents the network from fitting noise, which turns out to be the essential factor when the size of training samples is small.

Sprott's behavior illustrates a subtler point: the same attractor can produce sequences with very different entropy values depending on how they are discretized, while the classification performance does not differ much in each case. This means that ML fingerprinting exploits properties that are complementary to what standard randomness metrics quantify. Validating a single quantization use case can offer a false sense of security in face of potential attacks.

There are limitations of this study and some of them include the use of fixed parameter values for each system, the supervised setting in which the adversary assumes a predefined/known set of candidate attractors, and the absence of hardware-based validation. Open questions can focus on sensitivity to parameter variations and generalization to previously less studied attractors (such as Chen or Chua [15–18]).

6 Conclusions

We have shown that chaos-based cryptographic keystreams can be vulnerable to ML-based fingerprinting. A well-regularized ResNet model identifies the source attractor/chaotic system with high accuracy. This is based on short quantized

sequences, and the performance is above the performance achieved by simpler baseline models, including the Dense Simple and CNN Light architectures. High entropy values of the quantized chaotic sequences do not prevent this: systems that pass standard randomness tests can be easily classified because deep learning exploits structural signatures that standard tests do not measure and do not fully take into account.

Regularization is the decisive factor when we are faced with scarce data. Two architecturally identical networks analyzed in this study differ by nearly 40% points in accuracy on scarce binary data, with the gap explained mainly by the training strategy.

The discretization approach has an influence as well. The same attractor can generate sequences with very different entropy values depending on the quantization method, and remain equally identifiable by a classifier. As a consequence, evaluating only one quantization scheme can be misleading, and multiple schemes should be considered before drawing any conclusions and proceeding to practical deployment.

From a practical point of view: chaos-based cryptographic systems require evaluation against ML fingerprinting in addition to statistical tests. Chaotic keystream generators must use cryptographic post-processing such as SHA-256 [19] to destroy structural signatures; and new metrics quantifying resistance to ML classification should come side by side with standard entropy assessments.

The fingerprinting approach proposed in this work could also be applied beyond cryptography, for instance to classify chaotic dynamics in areas such as neuroscience, climate modeling, industrial IoT, and financial time-series analysis. Future work could focus on robustness to parameter changes and evaluation across multiple quantization schemes.

Acknowledgments. This work was supported by UEFISCDI and National University of Science and Technology Politehnica Bucharest through projects PN-IV-P7-7.1-PTE-2024-0505 and UNSTPB-industry collaboration project 34808/19.09.2025.

References

1. Dinu, A.: Singularity, observability, and independence: unveiling Lorenz's cryptographic potential. Mathematics **12**(18), 2798 (2024). https://doi.org/10.3390/math12182798
2. Pecora, L.M., Carroll, T.L.: Synchronization in chaotic systems. Phys. Rev. Lett. **64**(8), 821–824 (1990). https://doi.org/10.1103/PhysRevLett.64.821
3. Kocarev, L.: Chaos-based cryptography: a brief overview. IEEE Circuits Syst. Mag. **1**(3), 6–21 (2001). https://doi.org/10.1109/7384.963463
4. Dinu, A., Frunzete, M.: Singularity, observability and statistical independence in the context of chaotic systems. Mathematics **11**(2), 305 (2023). https://doi.org/10.3390/math11020305
5. Bassham, L., et al.: A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. NIST Special Publication 800-22. National Institute of Standards and Technology, Gaithersburg, MD (2010). https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=906762

6. Alvarez, G., Li, S.: Some basic cryptographic requirements for chaos-based cryptosystems. Int. J. Bifurcation Chaos **16**(8), 2129–2151 (2006)
7. Dinu, A.: From chaos to security: a comparative study of Lorenz and Rössler systems in cryptography. Cryptography **9**(3), 58 (2025). https://doi.org/10.3390/cryptography9030058
8. Dinu, A., Frunzete, M.: Image encryption using chaotic maps: development, application, and analysis. Mathematics **13**(16), 2588 (2025). https://doi.org/10.3390/math13162588
9. Sprott, J.C.: A dynamical system with a strange attractor and invariant tori. Phys. Lett. A **378**(20), 1361–1363 (2014). https://doi.org/10.1016/j.physleta.2014.03.028
10. Dinu, A., Vlad, A.: Romanian printed language, statistical independence and the type II statistical error. In: Proceedings of the 2021 International Conference on Speech Technology and Human-Computer Dialogue (SpeD), pp. 120–125 (2021). https://doi.org/10.1109/SpeD53181.2021.9587397
11. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, pp. 770–778 (2016). https://doi.org/10.1109/CVPR.2016.90
12. Fawaz, H.I., Forestier, G., Weber, J., Idoumghar, L., Muller, P.-A.: Deep learning for time series classification: a review. Data Min. Knowl. Discov. **33**(4), 917–963 (2019). https://doi.org/10.48550/arXiv.1809.04356
13. Maghrebi, H., Portigliatti, T., Prouff, E.: Breaking Cryptographic Implementations Using Deep Learning Techniques. In: Carlet, C., Hasan, M.A., Saraswat, V. (eds.) SPACE 2016. LNCS, vol. 10076, pp. 3–26. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-49445-6_1
14. Picek, S., Perin, G., Mariot, L., Wu, L., Batina, L.: SOK: deep learning-based physical side-channel analysis. ACM Comput. Surv. **55**(11), 227:1–227:35 (2023). https://doi.org/10.1145/3569577
15. Chen, M., Yu, J., Yu, Q., Li, C., Bao, B.: A memristive diode bridge-based canonical Chua's circuit. Entropy **16**, 6464–6476 (2014). https://doi.org/10.3390/e16126464
16. Yang, Z., et al.: A Chua's chaotic chirp spread-spectrum power spectral homogenization strategy based on distribution transformation. Electronics **13**(12), 2296. https://doi.org/10.3390/electronics13122296
17. Duan, Z., et al.: Integrated circuit of a Chua's system based on the integral-differential nonlinear resistance with multi-path voltage-controlled oscillator. Micromachines **15**(3), 401. https://doi.org/10.3390/mi15030401
18. Boobalan, S., Gurunathan Arthanari, S.K.: Lorenz and Chua chaotic key-based dynamic substitution box for efficient image encryption. Symmetry **17**(8), 1296. https://doi.org/10.3390/sym17081296
19. National Institute of Standards and Technology: Secure Hash Standard (SHS). FIPS Publication 180-4. NIST, Gaithersburg, MD (2015). https://doi.org/10.6028/NIST.FIPS.180-4

Physics-Informed Neural Networks for High-Dimensional Black–Scholes Option Pricing: Accuracy, Scalability, and Limitations

André Paz Gomes[1(✉)], Isaac P. Santos[1,2], and Lucia Catabriga[1]

[1] PPGI, Universidade Federal do Espírito Santo, Vitória, ES, Brazil
andre.p.gomes@edu.ufes.br, isaac.santos@ufes.br, luciac@inf.ufes.br
[2] DMA, Universidade Federal do Espírito Santo, São Mateus, ES, Brazil

Abstract. This work investigates the use of Physics-Informed Neural Networks (PINNs) for pricing European options under the Black-Scholes equation, including arithmetic basket options, and compares their performance with stochastic and classical numerical methods, namely Monte Carlo, finite differences, and binomial trees. The analysis spans multiple levels of dimensionality, enabling direct comparisons in one and two dimensions and assessing scalability in three and five dimensions. In low-dimensional cases, classical methods remain the most competitive in terms of accuracy and computational efficiency. Finite differences and binomial trees achieve the lowest errors, while Monte Carlo provides robust performance at a higher per-query cost. In these regimes, PINNs yield consistent approximations but do not outperform traditional methods under the adopted computational budget. In higher dimensions, PINNs demonstrate their main advantage by providing mesh-free approximations without the combinatorial growth associated with grid-based methods. Although they exhibit higher errors than Monte Carlo in the tested configurations, they remain applicable and scalable in regimes where deterministic methods become impractical. An additional advantage of PINNs lies in their offline-online structure, enabling fast and reusable inference, as well as efficient computation of sensitivities (Greeks) via automatic differentiation. However, their accuracy remains strongly dependent on training-related factors, with noticeable approximation bias in higher dimensions and near payoff nonsmoothness. Overall, PINNs should not be regarded as replacements for classical methods in low-dimensional problems, but rather as a promising alternative for high-dimensional option pricing, where scalability and flexibility are key requirements.

Keywords: Option pricing · Black–Scholes equation · Physics-informed neural networks · Monte Carlo methods · Finite difference methods · Binomial tree methods

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 313–329, 2027.
https://doi.org/10.1007/978-3-032-30530-5_20

1 Introduction

The pricing of financial derivatives relies on accurate and efficient numerical methods, particularly when no closed-form formula is available for the payoff of interest [6]. In the one-dimensional case of the Black-Scholes equation, standard European options admit an analytical solution [3]. However, multi-asset options, such as *basket* options, typically do not admit closed-form solutions for arithmetic payoffs and therefore require the use of numerical methods.

Among the classical approaches, Monte Carlo simulation, tree-based methods (binomial or multinomial), and finite difference discretizations are widely used [6]. Monte Carlo methods scale relatively well with dimensionality, but exhibit slow convergence and may require variance reduction techniques to improve efficiency. Finite differences and tree-based methods, in turn, are highly competitive for low-dimensional problems, but suffer from the curse of dimensionality. In grid-based discretizations, the number of degrees of freedom grows exponentially with the number of underlying assets, which quickly renders these methods impractical as dimensionality increases.

Physics-Informed Neural Networks (PINNs) have emerged as an alternative within the field of scientific machine learning, where neural networks are trained to satisfy the governing partial differential equations and the associated boundary and initial conditions through the minimization of a loss function [11]. In contrast to traditional grid-based numerical methods, PINNs approximate the solution using neural networks and therefore do not require the construction of structured meshes in high-dimensional spaces. This characteristic makes them an appealing candidate for multi-asset option pricing problems.

Recent studies have investigated the use of PINNs in financial applications, including option pricing problems in two dimensions [13] and more complex settings, such as Asian options, which are path-dependent [9]. In addition, the literature on learning-based methods for high-dimensional problems highlights the persistent challenge posed by the curse of dimensionality in option pricing and the ongoing search for scalable alternatives [15].

Despite these advances, a systematic comparison between PINNs and classical numerical methods for option pricing across multiple dimensions remains limited in the literature. In particular, the relative accuracy and computational cost of PINNs with respect to established approaches such as Monte Carlo simulation, finite difference schemes, and tree-based methods have not yet been thoroughly assessed in a unified experimental framework.

In this context, the objective of this work is to investigate whether Physics-Informed Neural Networks (PINNs) can provide a competitive alternative to classical numerical and stochastic approaches for option pricing problems of increasing dimensionality. To this end, we perform a consistent experimental comparison, in terms of MAE, RMSE, and total execution time, among PINNs, Monte Carlo simulation, a Crank–Nicolson finite difference scheme, and binomial tree methods in the one-dimensional case, where the analytical solution of the Black–Scholes equation is available as reference. The four methods are also compared in the two-dimensional case for a European *basket* option, using

a high-accuracy Monte Carlo estimate as the benchmark. Finally, a scalability assessment is conducted for three- and five-dimensional problems, comparing PINNs and Monte Carlo methods in order to reflect the practical infeasibility of tree-based and finite difference approaches in these regimes.

The remainder of this paper is organized as follows. Section 2 presents the Black-Scholes model and the considered option pricing problem. Section 3 describes the numerical methods, including the PINN formulation. Section 4 presents the experimental setup and results. Finally, Sect. 5 summarizes the conclusions.

2 The Black-Scholes Model in d Dimensions

Under the classical assumptions of the Black-Scholes model - a frictionless and arbitrage-free market, continuous trading, a constant risk-free rate r, and constant volatilities - consider a vector of d underlying assets $\mathbf{S}_t = (S_t^{(1)}, \ldots, S_t^{(d)})^\top$ each paying a continuous dividend yield (or carrying costs) q_i.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{Q})$ be a filtered probability space and let $\mathbb{Q}$ denote the risk-neutral measure. Under $\mathbb{Q}$, the asset prices follow d-dimensional correlated geometric Brownian motions [8]:

$$\frac{\mathrm{d}S_t^{(i)}}{S_t^{(i)}} = (r - q_i)\,\mathrm{d}t + \sigma_i\,\mathrm{d}W_t^{(i)}, \qquad i = 1, \ldots, d, \tag{1}$$

where $\sigma_i > 0$ denotes the volatility of asset i. The vector $\boldsymbol{W}_t = (W_t^{(1)}, \ldots, W_t^{(d)})^\top$ represents a d-dimensional Brownian motion whose components exhibit instantaneous correlations

$$\mathrm{d}W_t^{(i)}\,\mathrm{d}W_t^{(j)} = \rho_{ij}\,\mathrm{d}t, \qquad \rho_{ii} = 1, \qquad \boldsymbol{\rho} = [\rho_{ij}] \succeq 0, \tag{2}$$

where $\boldsymbol{\rho}$ is a symmetric positive semidefinite correlation matrix, ensuring a valid covariance structure for the driving noise processes [2].

Consider a European option with payoff $g(\mathbf{S}_T)$ at maturity T. Under the risk-neutral valuation principle, its arbitrage-free price is given by the conditional expectation

$$V(t, \mathbf{S}) = \mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)}\, g(\mathbf{S}_T) \,\middle|\, \mathbf{S}_t = \mathbf{S}\right], \tag{3}$$

which corresponds to the probabilistic representation of the option value under the risk-neutral measure [12].

By applying Itô's lemma to $V(t, \mathbf{S}_t)$ and invoking the no-arbitrage argument for self-financing portfolios, one obtains that V satisfies the multidimensional parabolic Black–Scholes partial differential equation:

$$\frac{\partial V}{\partial t} + \sum_{i=1}^{d}(r - q_i)S^{(i)}\frac{\partial V}{\partial S^{(i)}} + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\rho_{ij}\sigma_i\sigma_j S^{(i)}S^{(j)}\frac{\partial^2 V}{\partial S^{(i)}\partial S^{(j)}} - rV = 0, \tag{4}$$

subject to the terminal condition

$$V(T, \mathbf{S}) = g(\mathbf{S}). \tag{5}$$

The diffusive operator in (4) involves the covariance structure given by the matrix $(\rho_{ij}\sigma_i\sigma_j)$ and generates cross-derivative terms whenever correlations between assets are present. The indices i and j in the double summation run over all underlying assets, that is, $i, j = 1, \ldots, d$. Terms with $i = j$ correspond to the usual second-order derivatives with respect to each individual asset price, whereas terms with $i \neq j$ correspond to mixed second-order derivatives between different asset prices. These mixed derivatives arise from the instantaneous correlations ρ_{ij} between the Brownian motions driving the assets. Therefore, the diffusive operator in (4) involves the covariance structure given by the matrix $(\rho_{ij}\sigma_i\sigma_j)$ and generates cross-derivative terms whenever correlations between assets are present. These mixed derivatives significantly increase the numerical complexity of mesh-based discretizations, particularly finite difference schemes in two and higher spatial dimensions [14].

2.1 European Basket Option Payoff

We consider an arithmetic basket call option with payoff

$$g(\mathbf{S}_T) = \max\left\{\sum_{i=1}^{d} w_i S_T^{(i)} - K,\, 0\right\}, \qquad \sum_{i=1}^{d} w_i = 1,\ w_i \geq 0. \tag{6}$$

For $d = 1$ and $w_1 = 1$, this reduces to the standard European call option, for which a closed-form analytical solution is available [3]. For $d > 1$, the arithmetic basket option does not admit a known closed-form analytical solution, which motivates the use of numerical pricing methods [14].

2.2 Analytical Solution in 1D

In the one-dimensional case ($d = 1$), with European call payoff $g(S_T) = \max(S_T - K, 0)$, the Black-Scholes model admits a closed-form analytical solution. Under the risk-neutral measure $\mathbb{Q}$, the asset price follows a geometric Brownian motion, implying that $\ln(S_T)$ is normally distributed. This property allows the option price to be expressed as the expectation of the discounted payoff and leads to an explicit formula involving the cumulative distribution function of the standard normal distribution [8].

The Black-Scholes price of a European call option is given by [3,8]

$$V(S, t) = S\, e^{-q(T-t)}\Phi(d_1) - K\, e^{-r(T-t)}\Phi(d_2), \tag{7}$$

where q is the continuous dividend yield (when $q = 0$, the classical non-dividend-paying case is recovered),

$$d_1 = \frac{\ln(S/K) + (r - q + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \qquad d_2 = d_1 - \sigma\sqrt{T-t}, \tag{8}$$

and $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution.

The quantities $\Phi(d_1)$ and $\Phi(d_2)$ admit useful financial interpretations. In particular, $\Phi(d_2)$ corresponds to the risk-neutral probability that the option expires in the money, that is, $\mathbb{Q}(S_T > K)$. The term $e^{-q(T-t)}\Phi(d_1)$ appears in the expression for the option's *delta*, which measures the sensitivity of the price with respect to changes in the underlying asset [6].

In this work, Eq. (7) is used as the reference solution in the 1D scenario to measure the numerical error of Monte Carlo, binomial tree, finite differences, and PINN approaches.

3 Numerical Methods

Monte Carlo estimates the risk-neutral price of the option by simulating a large number of trajectories for the underlying assets under the measure $\mathbb{Q}$, followed by the computation of the sample average of the discounted payoff. For each simulated path, the asset dynamics follow the stochastic differential equation defined in Eq. (1). The option price is then approximated by

$$V(t, \mathbf{S}) \approx e^{-r(T-t)} \frac{1}{N} \sum_{k=1}^{N} g(\mathbf{S}_T^{(k)}),$$

where N denotes the number of simulated trajectories. Its accuracy improves as the number of paths increases, with statistical error decaying slowly, typically on the order of $1/\sqrt{N}$, which makes the method particularly attractive in higher dimensions due to its good scalability [5].

Tree-based methods approximate the stochastic dynamics of the underlying assets by a discrete-time recombining tree. In the one-dimensional case, the Cox-Ross-Rubinstein (CRR) binomial scheme is used. At each time step, the asset price may move upward or downward by predetermined factors, and the option value is obtained by backward propagation from the terminal payoff, discounting risk-neutral expectations step by step [4,6].

In the two-dimensional case, a correlated extension of the binomial tree is considered, in which joint probabilities are introduced to capture the correlation between the two underlying assets. However, the number of nodes grows rapidly with dimensionality, which makes tree-based methods computationally expensive in higher-dimensional settings.

Finite differences approximate the derivatives of the Black–Scholes PDE on a computational grid defined over the asset price variables and time, thereby transforming the continuous problem into a linear system of algebraic equations that is solved backward in time. In the one-dimensional case, the Crank–Nicolson scheme is applied. This implicit method is second-order accurate in both time and space and provides a good compromise between stability and accuracy when combined with appropriate boundary conditions [14].

In the two-dimensional case, an ADI (Alternating Direction Implicit) type scheme is used to reduce the computational cost associated with multidimensional grids. The domain is truncated and approximate boundary conditions are imposed to ensure numerical stability. However, in higher dimensions (3D and 5D), direct grid-based discretizations become impractical due to the exponential growth in the number of grid points and the associated memory and computational cost. For this reason, finite difference methods are not considered in these regimes [14].

3.1 PINN Formulation

Physics-Informed Neural Networks (PINNs) approximate the solution of a PDE by means of a neural network $u_\theta(t, \mathbf{S})$, parameterized by θ. The network is trained to satisfy simultaneously (i) the differential equation and (ii) the problem conditions, including the terminal condition and the boundary conditions, by incorporating such constraints directly into the loss function [11]. The main conceptual advantage of this approach is its mesh-free character: instead of explicitly discretizing the domain on a regular grid, as in finite differences, points are sampled in the continuous domain and violations of the PDE at those points are penalized during training.

Given the multidimensional Black-Scholes equation defined in Eq. (4), the differential operator $\mathcal{N}[\cdot]$ is defined as

$$\mathcal{N}[V](t, \mathbf{S}) = \frac{\partial V}{\partial t} + \sum_{i=1}^{d} (r - q_i) S^{(i)} \frac{\partial V}{\partial S^{(i)}} + \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \rho_{ij} \sigma_i \sigma_j S^{(i)} S^{(j)} \frac{\partial^2 V}{\partial S^{(i)} \partial S^{(j)}} - rV. \tag{9}$$

The PINN seeks a function u_θ such that $\mathcal{N}[u_\theta](t, \mathbf{S}) \approx 0$ for $(t, \mathbf{S})$ in the interior of the domain. For this purpose, the residual function is defined as

$$f_\theta(t, \mathbf{S}) = \mathcal{N}[u_\theta](t, \mathbf{S}), \tag{10}$$

computed through automatic differentiation. This technique enables the consistent evaluation of first- and second-order derivatives of the neural network output with respect to the inputs $(t, \mathbf{S})$, avoiding the need for finite difference approximations [11].

For European options, the terminal condition (payoff) is imposed at $t = T$:

$$u_\theta(T, \mathbf{S}) \approx g(\mathbf{S}) \equiv \max\left\{ \sum_{i=1}^{d} w_i S^{(i)} - K,\, 0 \right\}. \tag{11}$$

In addition, since the domain is truncated in practice, the asset prices are restricted to $S^{(i)} \in [0, S^{(i)}_{\max}]$. Approximate boundary conditions are imposed on the truncated domain. In particular, in the case of a call, low values of $S^{(i)}$ drive the price toward values close to zero, whereas high values make the price grow approximately linearly with the basket value, due to the dominance of the

payoff. In practice, boundary constraints consistent with the truncated computational domain are incorporated into the loss function.

Training is performed by minimizing a composite loss function that enforces the PDE residual, the terminal condition, and the boundary conditions:

$$\begin{aligned}\mathcal{L}(\theta) = \lambda_{\text{pde}} \frac{1}{N_f} \sum_{k=1}^{N_f} |f_\theta(\tau_k, \mathbf{S}_k)|^2 \\ + \lambda_{\text{ic}} \frac{1}{N_0} \sum_{m=1}^{N_0} |u_\theta(0, \mathbf{S}_m) - g(\mathbf{S}_m)|^2 \\ + \lambda_{\text{bc}} \frac{1}{N_b} \sum_{n=1}^{N_b} |u_\theta(\tau_n, \mathbf{S}_n) - b(\tau_n, \mathbf{S}_n)|^2 .\end{aligned} \tag{12}$$

where $\{(\tau_k, \mathbf{S}_k)\}$ denote collocation points sampled in the interior of the space-time domain, $\{(0, \mathbf{S}_m)\}$ correspond to points enforcing the payoff condition at $\tau = 0$, and $\{(\tau_n, \mathbf{S}_n)\}$ represent boundary points. The term with N_0 is referred to as the initial-condition loss because the original terminal condition $V(T, \mathbf{S}) = g(\mathbf{S})$ becomes $V(0, \mathbf{S}) = g(\mathbf{S})$ after introducing the time-to-maturity variable $\tau = T - t$. The weights $\lambda_{(\cdot)}$ control the balance between enforcing the PDE and imposing the terminal and boundary conditions, and adjustments are often required to ensure stability and good accuracy, especially in problems with different magnitude scales among the terms [11]. All loss components in Eq. (12) were assigned unit weights, i.e., $\lambda_{\text{pde}} = \lambda_{\text{ic}} = \lambda_{\text{bc}} = 1$, including each boundary-condition term in the multidimensional cases.

Unlike grid-based methods, the performance of PINN depends critically on the sampling of collocation points. Random sampling strategies, such as uniform sampling or Latin Hypercube sampling, are usually employed, and the density of points may be increased in regions where the solution is more difficult to approximate, such as near payoff nonsmoothness. Optimization is performed by gradient-based methods such as Adam and may be refined by quasi-Newton optimizers, such as L-BFGS, to accelerate convergence and reduce the final residual [11].

4 Numerical Experiments

We solve the Black-Scholes PDE introduced in Sect. 2 and evaluate the accuracy of the numerical solutions on a test set within the computational domain. The metrics considered are the mean absolute error (MAE) and the root mean square error (RMSE), defined by:

$$\text{MAE} = \frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} \left| V_i^{\text{num}} - V_i^{\text{ref}} \right|, \text{RMSE} = \left(\frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} \left(V_i^{\text{num}} - V_i^{\text{ref}} \right)^2 \right)^{1/2}, \tag{13}$$

where N_{test} is the evaluation points, V_i^{num} denotes the numerical approximation produced by a given method and V_i^{ref} denotes the corresponding reference value. Lower values of MAE and RMSE indicate better agreement with the reference solution. Computational cost is reported as the total execution time, including training for learning-based methods and evaluation on the test set.

In this work, the PINNs were implemented using the DeepXDE library with the `tensorflow.compat.v1` backend, which provides: (i) declarative definition of the PDE and the conditions, (ii) generation of collocation, boundary, and terminal points, (iii) automatic differentiation, and (iv) training and evaluation routines [7]. This enables a uniform implementation across different dimensions ($d = 1, 2, 3, 5$), preserving the same general loss formulation and changing only the dimension of the input vector and the corresponding differential operator.

The PINN is modeled as a fully connected feedforward neural network (FNN) that approximates the option price $V(S, \tau)$ on the computational domain $\Omega = [0, S_{\max}] \times [0, T]$, where $\tau = T - t$ denotes the time to maturity. The network takes (S, τ) as input and outputs a scalar approximation of V, the option value. The architecture $[2, 64, 64, 64, 1]$ is adopted, consisting of three hidden layers with 64 neurons each, using tanh activation and Glorot normal initialization. The choice of three hidden layers with 64 neurons each was motivated by a compromise between approximation capacity and computational cost. Since the Black–Scholes equation involves first- and second-order derivatives of the network output, excessively shallow networks may lack sufficient expressive power to represent the option price surface and its derivatives accurately. Conversely, substantially larger networks increase the cost of automatic differentiation and training, especially in higher-dimensional settings. Therefore, the architecture $[d+1, 64, 64, 64, 1]$ was adopted as a moderate-size fully connected network, which is consistent with moderate-size architectures previously used in PINN applications [1]. The same architectural pattern was maintained across all dimensions in order to ensure a consistent comparison framework. Input normalization is performed through feature scaling, with S normalized by $S_{\max}$ and τ by T, to improve numerical stability and facilitate optimization.

Training minimizes a composite loss function composed of the PDE residual evaluated at interior collocation points, together with constraints imposed by the initial condition in the transformed time variable, $V(S, 0) = \max(S - K, 0)$, and Dirichlet boundary conditions $V(0, \tau) = 0$ and the asymptotic condition $V(S_{\max}, \tau) \approx S_{\max} - Ke^{-r\tau}$. The training dataset consists of $N_f = 6000$ interior points, $N_b = 1000$ boundary points, and $N_0 = 1000$ initial points, automatically generated by the framework. Optimization is performed using the Adam method with learning rate 10^{-3} for 50,000 iterations.

All experiments were conducted on a machine with an Intel Xeon CPU (2.20 GHz), 12.7 GB RAM, and an NVIDIA Tesla T4 GPU (16 GB VRAM), running Python 3.10 and DeepXDE 1.6.1. This configuration was maintained consistently for all experiments to ensure fair comparisons of execution times.

Four experimental scenarios are considered, varying the dimensionality and the payoff structure. In all cases, maturity $T = 1$ and risk-free rate $r = 0.05$ are

fixed. For each scenario, we define: (i) a test set used exclusively for computing MAE and RMSE, and (ii) a reference solution, obtained analytically when available or via high-accuracy Monte Carlo simulations otherwise.

In the 1D case, a European call option is considered with $S_{\max} = 200$, $K = 50$, $T = 1$, $r = 0.05$, and volatility $\sigma = 0.20$. Evaluation is performed on a regular 40×40 grid over the computational domain, Ω. The analytical Black-Scholes formula is used as reference, providing a deterministic ground truth and eliminating statistical uncertainty in the error metrics [3].

In higher dimensions, European basket options are considered with nonnegative and normalized weights ($\sum_i w_i = 1$), strike $K = 50$, $T = 1$, and $r = 0.05$. In the 2D case, equal weights $w_1 = w_2 = 0.5$, volatilities $(\sigma_1, \sigma_2) = (0.20, 0.25)$, and correlation $\rho = 0.30$ are adopted. Since no closed-form solution exists, the reference is computed via Monte Carlo with 600,000 paths. The test set consists of 400 randomly sampled points.

Similarly, in the 3D case, equal weights $w_1 = w_2 = w_3 = \frac{1}{3}$, volatilities $\sigma = (0.20, 0.25, 0.22)$, and correlations $(\rho_{12}, \rho_{13}, \rho_{23}) = (0.30, 0.20, 0.25)$ are used, with the same Monte Carlo configuration and test set size. Finally, in the 5D case, equal weights $w_i = \frac{1}{5}$, $i = 1, 2, \cdots, 5$, volatilities $\sigma = (0.20, 0.25, 0.22, 0.18, 0.24)$, and uniform correlation $\rho_{ij} = 0.25$, for $i \neq j$, are considered. The reference solution is again obtained via Monte Carlo with 600,000 paths, and the test set contains 400 random points. This last scenario highlights the regime in which deterministic grid-based methods become impractical, emphasizing the scalability of PINNs and Monte Carlo methods under a common evaluation protocol.

4.1 1D Case: Four Methods Versus Analytical Solution

In the one-dimensional case, the comparison is performed against the closed-form Black-Scholes formula for the European call option, allowing the numerical error to be assessed directly throughout the computational domain. To mitigate oscillations and numerical noise induced by the nonsmooth payoff near $\tau \approx 0$, the finite difference scheme uses the Crank-Nicolson (CN) method with Rannacher smoothing, consisting of two half-steps of implicit Euler prior to CN, following the original proposal [10] and its well-established use in option pricing with nonsmooth payoffs.

PINN training in this experiment required 161.765 s, a cost that was incurred only during the offline phase. After training, the network can be reused for multiple price and Greeks queries with reduced inference cost. Similarly, the finite difference method requires a PDE solution stage (0.499 s) to construct the surface on the grid; subsequent queries are dominated by interpolation over the precomputed grid. Consequently, Table 1 reports only the prediction time on the test set (*pred_time*), highlighting the cost per query after any offline computation.

Table 1 highlights the trade-off between accuracy and cost per query. The binomial tree achieves the smallest global error, but exhibits the highest prediction time, since each query involves a backward evaluation over the tree structure. With Rannacher smoothing, the finite difference method combines

Table 1. 1D case (European call). Errors on the test set with respect to the analytical price and prediction time (*pred_time*) on the test set.

Method	MAE	RMSE	*pred_time* [s]
Finite Difference	3.67×10^{-4}	1.07×10^{-3}	$\mathbf{2.66 \times 10^{-3}}$
Binomial Tree	$\mathbf{9.50 \times 10^{-5}}$	$\mathbf{3.12 \times 10^{-4}}$	7.01
Monte Carlo	7.25×10^{-3}	1.21×10^{-2}	6.13
PINN	1.39×10^{-1}	1.57×10^{-1}	2.22×10^{-1}

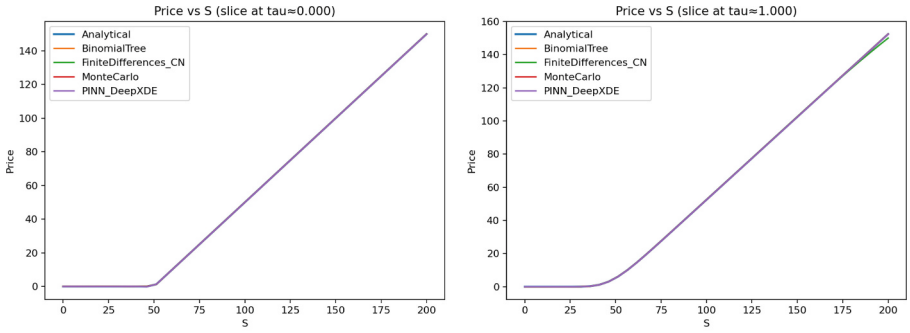
extremely fast prediction, via interpolation over the precomputed surface, with low aggregate error, rendering it competitive with the binomial tree in terms of accuracy while being significantly faster in prediction. Monte Carlo provides moderate accuracy, but at a high prediction cost, as each price requires a large number of simulated paths to reduce variance. The PINN, in this configuration, presents significantly higher error despite its low inference cost after training. This behavior reflects approximation bias and potentially the need for better sampling and better balancing of the loss weights under the adopted computational budget.

To complement the quantitative analysis, Fig. 1 presents price profiles $C(S, \tau)$ as functions of S at two representative stages of the contract: near maturity ($\tau \approx 0$) and at the initial time ($\tau \approx T$). A close agreement is observed between the analytical solution and the binomial tree and finite difference approximations. Near maturity ($\tau \approx 0$), all methods accurately capture the payoff structure, despite its nonsmoothness, due to the explicit inclusion of training points enforcing the initial condition. For larger values of τ, small deviations become more noticeable, as the solution is governed by the PDE dynamics and must be approximated through the learned representation.

Figure 2 summarizes the distribution of the absolute error across the test set on a logarithmic scale. The binomial tree exhibits the smallest errors, reaching values close to machine precision, although with a wider spread that indicates localized larger errors. The finite difference method exhibits slightly higher error with moderate dispersion, reflecting a more uniform accuracy across the domain, which can be attributed to the use of Rannacher smoothing. Monte Carlo shows consistently larger errors, but with a relatively narrow distribution, indicating a more uniform error pattern dominated by sampling variability, whereas the PINN displays higher errors throughout most of the domain, consistent with the aggregate MAE and RMSE values reported in Table 1.

4.2 2D Case: Four Methods (Monte Carlo Reference)

In the two-dimensional case, corresponding to a European basket option, no general closed-form formula exists to serve as a direct reference throughout the domain. Therefore, a high-accuracy Monte Carlo simulation with 600,000 paths is used as a reference, exclusively for error computation. In this experiment,



(a) Maturity: $\tau \approx 0.000$ (b) Initial time: $\tau \approx 1.000$

Fig. 1. 1D case. Profiles of $C(S, \tau)$ at maturity and at the initial time ($\tau = T - t$), comparing the analytical solution and the numerical approximations.

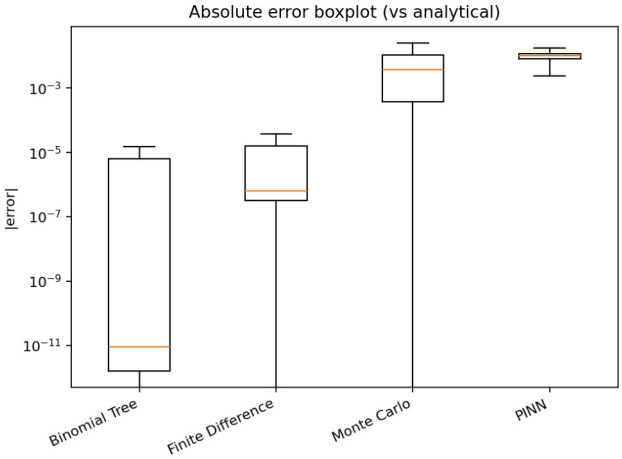


Fig. 2. 1D case. Boxplot of the absolute error (log scale) with respect to the analytical solution, aggregating errors over the test set.

the PINN was trained for 50,000 iterations with periodic resampling of PDE-points, resulting in 185.694 s of training in the offline phase. After this one-time cost, inference becomes significantly faster. Similarly, the two-dimensional finite difference solution was obtained using a fully implicit Crank-Nicolson scheme, requiring 0.136 s to construct the solution surface; subsequent predictions are dominated by interpolation. To emphasize the cost per query after offline stages, Table 2 reports the prediction time on the test set (*pred_ time*) along with aggregate errors relative to the reference.

To complement the aggregate results in Table 2, Fig. 3 presents the distribution of the absolute error over the test set in logarithmic scale, highlighting differences in dispersion and the presence of outliers.

Table 2. 2D case (European basket option). Errors with respect to the high-accuracy Monte Carlo reference and prediction time (*pred_ time*) on the test set.

Method	MAE	RMSE	*pred_ time* [s]
Finite Difference	5.08×10^{-3}	7.09×10^{-3}	$\mathbf{3.85 \times 10^{-4}}$
Binomial Tree	$\mathbf{4.19 \times 10^{-3}}$	$\mathbf{5.33 \times 10^{-3}}$	2.89×10^{02}
Monte Carlo	9.18×10^{-3}	1.19×10^{-2}	1.25
PINN	1.06×10^{-2}	1.95×10^{-2}	2.53×10^{-1}

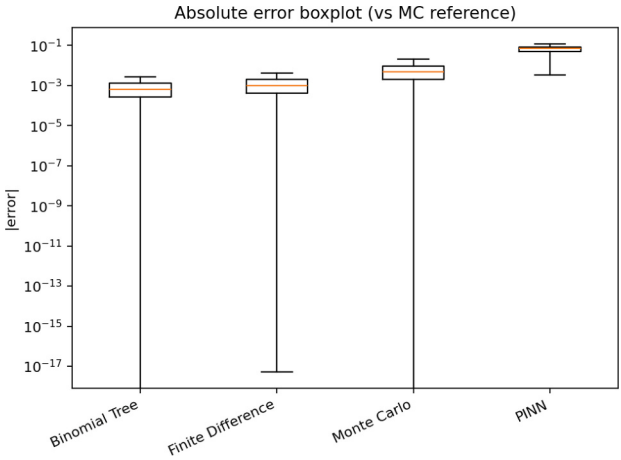


Fig. 3. 2D case. Boxplot of the absolute error (log scale) with respect to the Monte Carlo reference, aggregating errors over the test set.

Table 2 and Fig. 3 reveal four distinct performance regimes. The two-dimensional binomial tree, considering N=120, achieves the smallest aggregate error, but at an extremely high prediction time, reflecting the combinatorial growth in the number of nodes. The finite difference method with Crank–Nicolson exhibits low error and negligible *pred_ time*, as queries reduce to interpolation over the precomputed grid. Monte Carlo with 60,000 paths provides a robust balance between cost and accuracy, with error consistent with sampling variance and controllable through the number of paths. The PINN, while mesh-free and capable of fast inference after training, remains less accurate than Monte Carlo in this experiment, particularly in terms of RMSE, indicating the presence of approximation bias under the adopted training budget and sampling strategy.

4.3 3D Case: PINN Versus Monte Carlo

In the three-dimensional case, corresponding to a European basket option, a high-accuracy Monte Carlo with 600,000 paths is adopted as a reference solution. This reference is used exclusively to compute the errors of both the Monte Carlo

method with 80,000 paths and the PINN predictions. Table 3 summarizes the results for the two evaluated methods, reporting aggregate errors relative to this reference and the prediction time (*pred_time*) on the test set. For the PINN, it is important to highlight the existence of an offline training cost, which in this experiment amounts to 241.972955 s, while pointwise inference remains fast once training is completed.

Table 3. 3D case (European basket option). Errors with respect to the high-accuracy Monte Carlo reference and prediction time (*pred_time*) on the test set.

Method	MAE	RMSE	*pred_time* [s]
Monte Carlo	$\mathbf{3.89 \times 10^{-3}}$	$\mathbf{5.68 \times 10^{-3}}$	1.31
PINN	6.57×10^{-2}	2.06×10^{-1}	$\mathbf{3.95 \times 10^{-1}}$

In this three-dimensional setting, the analysis is restricted to Monte Carlo and PINNs, since deterministic grid-based methods rapidly become computationally prohibitive in terms of memory and execution time as dimensionality increases. Figure 4 illustrates the distribution of the pointwise absolute error over the test set in logarithmic scale.

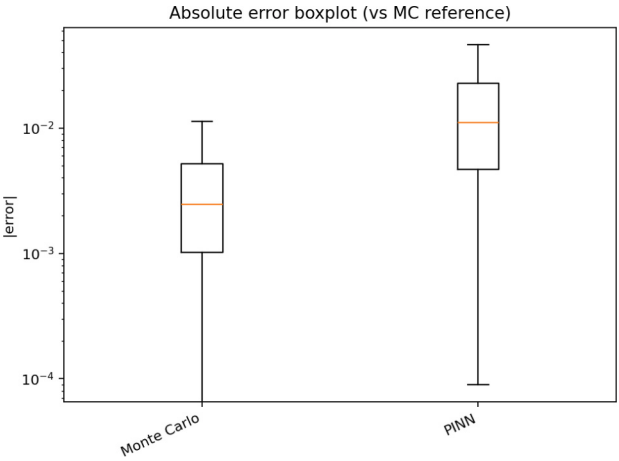


Fig. 4. 3D case. Boxplot of the absolute error (log scale) with respect to the Monte Carlo reference, aggregating errors over the test set.

Table 3 shows that Monte Carlo maintains low aggregate error at moderate computational cost under the adopted sampling budget, constituting a strong baseline in three dimensions when accuracy can be controlled via the number of simulated paths. The PINN, in contrast, retains the advantage of being mesh-free

and achieves faster inference on the test set, but at the expense of substantially higher aggregate error. Figure 4 reinforces this observation by showing that the absolute error distribution of the PINN is shifted toward higher values compared to Monte Carlo. This behavior suggests the predominance of approximation bias, in addition to variance effects, which is expected when learning solutions associated with a nonsmooth payoff. Overall, in the evaluated configuration, Monte Carlo remains superior in terms of accuracy, whereas the PINN provides a viable alternative in higher dimensions, characterized by a high offline cost but fast and reusable inference once trained. This result highlights the increasing difficulty of balancing approximation accuracy and computational cost for PINNs as dimensionality grows, reinforcing the need for improved training strategies in high-dimensional settings.

4.4 5D Case: PINN Versus Monte Carlo

In the five-dimensional case, corresponding to a European basket option, a high-accuracy Monte Carlo simulation with 600,000 paths is adopted as a reference solution. This reference is used exclusively to compute the errors of both the Monte Carlo method with 80,000 paths and the PINN predictions on the test set. Table 4 reports the aggregate errors, MAE and RMSE, relative to this reference, along with the prediction time (*pred_ time*) on the test set. For the PINN, the existence of an offline training cost must be emphasized, amounting in this experiment to 416.840 s, while inference remains fast once training is completed.

Table 4. 5D case (European basket option). Errors with respect to the high-accuracy Monte Carlo reference and prediction time (*pred_ time*) on the test set.

Method	MAE	RMSE	*pred_ time* [s]
Monte Carlo	$\mathbf{3.68 \times 10^{-3}}$	$\mathbf{4.86 \times 10^{-3}}$	1.90
PINN	2.57×10^{-2}	8.17×10^{-2}	$\mathbf{4.21 \times 10^{-1}}$

As in the three-dimensional case, Fig. 5 presents the distribution of the absolute error on the test set in logarithmic scale.

The five-dimensional results further highlight the role of PINNs as mesh-free approximators in regimes where grid-based discretizations become computationally prohibitive. Table 4 shows that Monte Carlo maintains an excellent trade-off between accuracy and computational cost under the adopted sampling budget, providing a strong and reliable baseline. The PINN, in contrast, remains applicable without suffering from the combinatorial explosion associated with grid-based methods and achieves faster inference on the test set. However, this comes at the expense of higher aggregate error and a computational cost dominated by the offline training phase. Figure 5 shows that the error distribution of the PINN is shifted toward higher values compared to Monte Carlo, indicating that performance in this configuration is primarily limited by approximation bias, as well

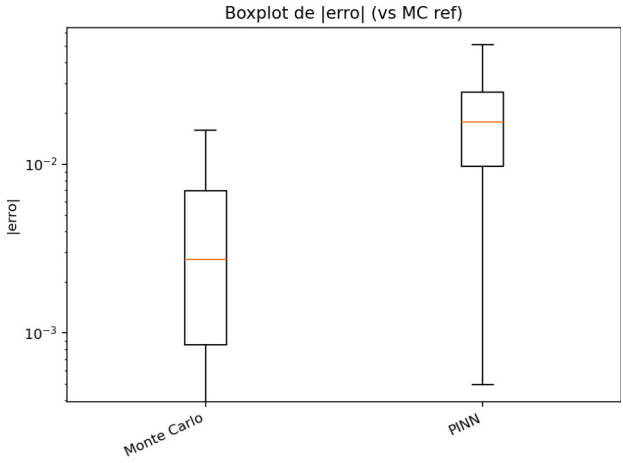


Fig. 5. 5D case. Boxplot of the absolute error (log scale) with respect to the Monte Carlo reference, aggregating errors over the test set.

as by the difficulty of simultaneously enforcing the PDE residual and the nonsmooth terminal condition in higher dimensions. Nevertheless, the PINN remains attractive as a reusable global surrogate, particularly in scenarios requiring multiple pricing evaluations and the efficient computation of Greeks after training.

5 Conclusion

This work investigated the use of Physics-Informed Neural Networks (PINNs) for pricing European options under the Black–Scholes equation, including arithmetic basket options, and compared their performance with classical numerical methods such as Monte Carlo, finite differences, and binomial trees. The analysis spanned multiple levels of dimensionality, enabling direct comparisons in 1D and 2D, as well as an assessment of scalability in 3D and 5D, where deterministic mesh-based methods become computationally prohibitive.

The results indicate that, in low-dimensions settings, classical methods remain the most competitive choices in terms of accuracy and computational cost. In 1D, the binomial trees and the finite differences achieved the smallest errors relatives to the analytical solution, while Monte Carlo provided consistent results at a higher per-query cost. In 2D, binomial trees again yielded the highest accuracy, but with prohibitive prediction time, whereas finite differences offered an excellent balance between accuracy and efficiency. In these regimes, PINNs proved to be functional and stable, but less competitive in accuracy under the adopted computational budget.

In contrast, the experiments in 3D and 5D highlight the main potential of PINNs: their applicability in high-dimensional settings, where grid-based methods suffer from the curse of dimensionality. Although PINNs exhibited higher

aggregate errors than Monte Carlo in the tested configurations, they maintained approximation capability without the combinatorial explosion associated with multidimensional discretization. This reinforces their relevance as a mesh-free alternative for high-dimensional problems.

An important advantage of PINNs is that the computational cost is largely concentrated in the offline phase. Once trained, the model can be reused for multiple pricing and sensitivity (Greeks) evaluations, with low inference cost. Moreover, the continuous nature of the approximation enables the direct computation of sensitivities via automatic differentiation, which is particularly attractive in applications involving repeated queries, calibration, or optimization.

At the same time, the experiments reveal important limitations. The PINN exhibited significant approximation bias, especially in higher dimensions and in regions influenced by payoff nonsmoothness. This behavior highlights the strong dependence of performance on factors such as sampling strategy, loss balancing, network architecture, variable scaling, and training procedures. Therefore, the results presented here should be interpreted as representative of a specific implementation, rather than as a general assessment of all PINN-based approaches.

Overall, PINNs should not be regarded as universal replacements for classical methods, particularly in low-dimensional problems, where finite differences and binomial trees remain superior. However, they emerge as a promising alternative for high-dimensional regimes, where scalability becomes a critical concern. Their ability to provide a reusable global approximation, combined with efficient inference and direct access to Greeks, makes them especially appealing when a continuous surrogate model is required.

Future work includes improving collocation sampling strategies, particularly near nonsmooth regions; developing adaptive loss-balancing techniques; exploring deeper or problem-specific network architectures; incorporating hard constraints to enforce boundary and terminal conditions; and extending the framework to more complex derivatives, such as American, Asian, and path-dependent options. In addition, hybrid approaches combining PINNs with Monte Carlo or other numerical methods represent a promising direction for leveraging both statistical robustness and approximation flexibility. These findings contribute to a clearer understanding of the practical role of PINNs in computational finance, particularly in relation to the trade-off between scalability and accuracy in high-dimensional problems.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT (OpenAI) in order to assist with language revision and text improvement. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Acknowledgments. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 - and

by the Fundação de Amparo à Pesquisa e Inovação do Espírito Santo (FAPES), Grant Notice N° 02/2026.

References

1. Arap, O., et al.: Gabor-enhanced physics-informed neural networks for fast simulations of acoustic wavefields. arXiv preprint arXiv:2502.17134 (2025). https://arxiv.org/abs/2502.17134
2. Björk, T.: Arbitrage Theory in Continuous Time, 3rd edn. Oxford University Press, Oxford (2009)
3. Black, F., Scholes, M.: The pricing of options and corporate liabilities. J. Polit. Econ. **81**(3), 637–654 (1973)
4. Cox, J.C., Ross, S.A., Rubinstein, M.: Option pricing: a simplified approach. J. Financ. Econ. **7**(3), 229–263 (1979)
5. Glasserman, P.: Monte Carlo Methods in Financial Engineering. Springer, New York (2003)
6. Hull, J.: Options, Futures, and Other Derivatives, 10th edn. Pearson, Harlow (2017)
7. Lu, L., Meng, X., Mao, Z., Karniadakis, G.E.: Deepxde: a deep learning library for solving differential equations. J. Comput. Phys. **429**, 109926 (2021)
8. Merton, R.C.: Theory of rational option pricing. Bell J. Econ. Manag. Sci. **4**(1), 141–183 (1973)
9. Park, S., Moon, K.S., Kim, H.: Asian option pricing using the physics-informed neural networks method. Econom. Comput. Econom. Cybernet. Stud. Res. **59**(1), 5–20 (2025). https://doi.org/10.24818/18423264/59.1.25.01
10. Pooley, D.M., Vetzal, K.R., Forsyth, P.A.: Convergence remedies for non-smooth payoffs in option pricing. University of Waterloo, Tech. rep. (2002)
11. Raissi, M., Perdikaris, P., Karniadakis, G.E.: Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J. Comput. Phys. **378**, 686–707 (2019)
12. Shreve, S.E.: Stochastic Calculus for Finance II: Continuous-Time Models. Springer, New York (2004)
13. Tavakoli, et al.: European option pricing with two underlying assets using Pinn (2024), publication details to be completed
14. Tavella, D., Randall, C.: Pricing Financial Instruments: The Finite Difference Method. Wiley, New York (2000)
15. Zhuang, J., Ding, D., Lu, W., Wu, X., Yuan, G.: A gaussian process based method with deep kernel learning for pricing high-dimensional American options (2024). arXiv preprint arXiv:2311.07211

6th Workshop on: Privacy in the Cloud/Edge/IoT World (PCEIoT 2026)

A Lightweight CRYSTALS-Dilithium Post-quantum Cryptography Signature for FPGA-Based Edge Computing

Huy Minh Dung Do[1,2], Duc Hoang Pham[1,2], Thien Hai Ha[1,2], and Cuong Pham-Quoc[1,2(✉)]

[1] Ho Chi Minh City University of Technology (HCMUT), Ho Chi Minh City, Vietnam
{dung.do05050,hoang.pham2004,hai.hathien,cuongpham}@hcmut.edu.vn
[2] Vietnam National University - Ho Chi Minh City (VNU-HCM), Ho Chi Minh City, Vietnam

Abstract. In this golden time of edge computing, while more IoT devices and distributed services are increasingly processing security-sensitive data close to the data source, efficient authentication has become a critical requirement under strict latency, power, and resource constraints. At the same time, advances in quantum computing threaten the widespread deployment of public-key cryptography, motivating the transition to post-quantum digital signatures, such as CRYSTALS-Dilithium, which has been standardized by NIST as ML-DSA (Module-Lattice-Based Digital Signature Algorithm). However, practical ML-DSA deployment on edge platforms remains challenging due to compute- and memory-intensive kernels, including NTT/INTT-based polynomial arithmetic, SHAKE-based hashing, and signing-time rejection sampling, as well as non-negligible data movement overhead in SoC systems. This paper presents a lightweight FPGA-based ML-DSA signature accelerator tailored for FPGA-based edge computing, following a HW/SW co-design approach that offloads arithmetic-dominant kernels to programmable logic while retaining protocol orchestration and standard-compliant serialization in software. We implement the proposed design on the Xilinx Kria KV260 edge platform and evaluate it in terms of hardware utilization, timing, signing and verification latency, throughput, and power/energy consumption. Experimental results demonstrate that the FPGA-accelerated design enhances end-to-end performance compared to a software-only baseline on the embedded ARM processor, while maintaining a compact resource footprint and energy efficiency suitable for power-sensitive edge deployments. The proposed architecture is modular and can be integrated with other edge security functions, enabling practical post-quantum authentication for IoT gateways and edge services.

Keywords: FPGA · Hardware accelerator · Post-quantum cryptography · ML-DSA · CRYSTALS-Dilithium · Edge Computing · IoT

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 333–350, 2027.
https://doi.org/10.1007/978-3-032-30530-5_21

1 Introduction

In the era of edge computing, as IoT systems and AI-enabled services continue to move computation closer to data sources [10,25], security and privacy have become critical concerns due to the reliance on distributed devices operating under strict resource and power constraints. At the same time, advances in quantum computing pose a fundamental threat to widely deployed public-key cryptography. In particular, Shor's algorithm makes RSA and elliptic-curve cryptography vulnerable once large-scale quantum computers become practical [21], raising serious concerns about long-lived data and the authentication mechanisms that underpin modern networks and services [8]. This challenge is amplified by the "harvest now, decrypt later" paradigm, where adversaries can collect protected communications today in anticipation of future decryption capabilities [1,15,16]. In response, NIST has standardized a first set of post-quantum cryptographic algorithms, including ML-DSA (CRYSTALS-Dilithium) for digital signatures [3,17].

Although ML-DSA provides a promising path toward quantum-resistant authentication, deploying it efficiently on edge platforms remains challenging. Dilithium relies on lattice-based hardness assumptions, primarily Module Learning With Errors and Module Short Integer Solution [12,19], which are widely regarded as resistant to classical and quantum adversaries [4,18]. However, practical implementations must execute compute and memory-intensive kernels, including NTT-based polynomial arithmetic, SHAKE-based hashing and rejection sampling, and must do so while meeting real-time constraints and tight energy budgets [17,26]. These requirements are particularly relevant in edge-centric security settings, where digital signatures are needed to provide integrity, authenticity, and non-repudiation [6], and where security workloads are often offloaded from constrained devices to nearby edge nodes (e.g., security agents) to improve scalability [9,20].

Field-Programmable Gate Arrays (FPGAs) offer an attractive solution to address these challenges by combining reconfigurability with energy-efficient parallelism. This makes FPGAs particularly suitable for edge platforms that must deliver predictable performance under tight power and resource budgets, while still retaining the flexibility to adapt to evolving cryptographic standards and deployment requirements [13]. In this paper, we present *A Lightweight CRYSTALS-Dilithium Post-Quantum Cryptography Signature for FPGA-based Edge Computing*. Instead of treating the edge node as a generic compute accelerator, we focus on the digital-signature workload itself and design a compact Dilithium (ML-DSA) signature engine tailored to edge constraints. The proposed design focuses on the practical sign and verification at the edge, targeting low latency and resource efficiency while preserving the security properties required by the standardized algorithm [17]. To this end, we focus on the dominant computations in Dilithium, including NTT-based polynomial arithmetic and SHAKE-based hashing, and expose a hardware/software interface that supports integration into edge applications and security offload settings [26].

The proposed FPGA-based signature engine is evaluated with various parameters, including hardware utilization, power efficiency, and signature performance (latency and throughput). The results indicate that our approach is well-suited for security-sensitive edge environments such as IoT networks, industrial monitoring, and real-time data acquisition systems, where quantum-resistant authentication must be delivered within limited resources [10].

The main contributions of our paper can be summarized in threefold.

1. We propose a lightweight FPGA-based CRYSTALS-Dilithium (ML-DSA) digital signature design for FPGA-based edge computing, focusing on resource-efficient and low-latency signing and verification.
2. We demonstrate the proposed design with a prototype on the Xilinx Kria KV260 edge computing platform and report end-to-end performance in a realistic SoC-FPGA setting.
3. We present and analyze experimental results in terms of hardware utilization and signature performance to facilitate reproducible comparisons and future optimization studies for post-quantum signature deployment at the edge.

The rest of the paper is structured as follows. Section 2 presents the background on the work related to CRYSTALS-Dilithium/ML-DSA and FPGA-based post-quantum systems. We introduce our proposed FPGA-based Dilithium signature design in Sect. 3. Section 4 discusses our prototype with Xilinx Kria KV260 edge computing. Experiments with the prototype platform and different configurations are analyzed in Sect. 5. Finally, Sect. 6 concludes our paper.

2 Preliminaries and Related Work

In this section, we first summarize the background of post-quantum digital signatures and CRYSTALS-Dilithium (standardized as ML-DSA in FIPS 204) that are relevant to FPGA-based edge deployments. We highlight the underlying hardness assumptions and the main computational kernels that dominate practical implementations. We then analyze related work on FPGA acceleration of ML-DSA, with emphasis on SoC-FPGA co-design considerations such as workload partitioning and data movement.

2.1 CRYSTALS-Dilithium (ML-DSA) Preliminaries

CRYSTALS-Dilithium is a lattice-based digital signature scheme selected in the NIST Post-Quantum Cryptography standardization effort and standardized as the Module-Lattice-Based Digital Signature Standard (ML-DSA) under FIPS204 [3,17]. Digital signatures remain a core mechanism for ensuring data authenticity, integrity, and non-repudiation in distributed systems, including edge and IoT environments where devices frequently exchange security-critical messages. FIPS 204 specifies three sets of standardized parameters, namely ML-DSA-44, ML-DSA-65 and ML-DSA-87, corresponding to increasing security strengths and implementation costs [17]. In many edge use cases that require

long-term trust in signed data, the highest-security parameter set is attractive, but it also imposes the heaviest arithmetic workload and the pressure of the memory bandwidth [17].

From a security point of view, ML-DSA targets strong unforgeability in chosen-message attacks, based on the hardness of the module-lattice assumptions [17]. In particular, FIPS 204 relates the security of ML-DSA to the hardness of Module Learning With Errors (MLWE) and a structured MSIS-type problem called SelfTargetMSIS. Informally, a successful forgery would imply an efficient solver for these underlying module-lattice problems [12,19]. The credibility of LWE-like assumptions is supported by worst-case to average-case reductions that connect average-case problems encountered by an attacker to worst-case lattice problems such as GapSVP and SIVP, with subsequent work strengthening the classical hardness picture for structured settings [4,12,18,19].

From an implementation perspective, ML-DSA is dominated by structured polynomial arithmetic. Computations are carried out on the polynomial ring $R_q = \mathbb{Z}_q[x]/(x^{256} + 1)$, and efficient polynomial multiplication is based on the Number-Theoretic Transform (NTT) and its inverse [17]. This NTT-centric structure appears repeatedly in key generation, signing, and verification, for example when computing commitments of the form $w = Ay$ using NTT-domain multiplication. In addition, the signing procedure contains a rejection sampling loop that repeatedly generates a signature candidate until norm constraints are met; this step is essential to prevent leakage of secret information through out-of-bound outputs [17]. ML-DSA also uses SHAKE128 and SHAKE256 extensively for matrix expansion, sampling, and challenge generation, and supports a pre-hash interface (HashML-DSA) for deployments that sign message digests rather than raw messages [17].

These characteristics motivate hardware acceleration on edge-class SoC-FPGA platforms, where compute resources and energy budgets are limited, but tight latency requirements are common [10,25]. A practical design pattern is HW/SW co-design: arithmetic-intensive, data-parallel kernels (NTT/INTT, pointwise multiplication, modular arithmetic) are mapped to the programmable logic, while control-heavy tasks (protocol orchestration, parameter checks, serialization, and rejection sampling decision logic) remain on the processing system. This split is typically paired with AXI-based integration, using memory-mapped control paths for configuration and DMA streaming for bulk polynomial data transfers, while on-chip BRAM acts as a low-latency scratchpad. To keep the design lightweight and timing-robust on FPGA fabric, stream-oriented interfaces with moderate data widths, for instance, transferring seeds or coefficients in 64-bit chunks, are often preferable to very wide buses that increase routing congestion and critical path length.

2.2 FPGA-Based Dilithium Related Work

FPGA-based acceleration of CRYSTALS-Dilithium has been actively studied because lattice-based signatures impose substantially higher arithmetic and memory costs than classical schemes, which makes software-only deployment

difficult under edge-class constraints. Across prior implementations, the dominant runtime is typically concentrated in structured polynomial arithmetic (especially NTT/INTT and pointwise multiplication) and SHAKE-based hashing, while system-level efficiency is often bounded by memory bandwidth and the overhead of moving polynomial data between the processing system and programmable logic. Consequently, most FPGA-oriented studies can be understood through three recurring design goals: increasing NTT throughput, reducing data-movement overhead, and selecting a practical HW/SW partition that preserves flexibility without sacrificing end-to-end performance.

Therefore, a large part of the literature focuses on NTT-centric optimization, since NTT/INTT is repeatedly invoked in key generation, signing, and verification. Beckwith *et al.* present a high-performance FPGA implementation of CRYSTALS-Dilithium and demonstrate that carefully engineered transform pipelines and polynomial arithmetic datapaths can achieve very low latency compared to earlier FPGA efforts [2]. In a similar spirit, Kundi *et al.* propose a high-performance NTT hardware accelerator intended to support both ML-KEM and ML-DSA, reflecting the broader trend of designing reusable NTT engines that can serve multiple NIST PQC primitives [11]. At the framework level, unified NTT multiplication architectures have also been explored to reduce duplicated hardware and improve portability across platforms, which is attractive in edge deployments where resource budgets are tight and multiple PQC services may coexist [13]. These works collectively indicate that the transform datapath is necessary but not sufficient: the realized system benefit depends on whether the design can keep the NTT pipeline continuously fed and can manage coefficient storage and reordering without introducing stalls.

Beyond arithmetic throughput, several studies show that data movement can become the primary bottleneck once polynomial kernels are accelerated. Carril *et al.* investigate hardware acceleration for high-volume operations of CRYSTALS-Kyber and CRYSTALS-Dilithium and emphasize an execution model that leverages batch processing to maximize effective memory throughput, together with hardware optimization of selected algorithmic components [5]. This direction highlights that host–accelerator interaction, buffering strategy, and transfer granularity may dominate end-to-end latency when the accelerator is invoked frequently and that throughput-oriented designs can benefit substantially from amortizing setup and transfer costs across many requests. While such batch-centric acceleration is well matched to high-volume services (for example, gateway-style nodes that process many signature requests), it also underlines a key challenge for edge settings: if the target is low latency for single signing or verification operations, the communication path and invocation overhead must be explicitly co-optimized with the compute core.

A third line of work addresses precisely this co-optimization through HW/SW partitioning on SoC-FPGA platforms. Mao *et al.* present a SW/HW configurable high-performance co-design of CRYSTALS-Dilithium, proposing a hardware accelerator integrated into a flexible SoC architecture and emphasizing a parameterized design that enables runtime configuration across Dilithium

modes [14]. Their results reinforce that accelerator effectiveness depends not only on fast NTT and hashing modules, but also on how control-dominant tasks and data marshaling are handled in software, and on how efficiently the PS–PL interface supports repeated movement of polynomial vectors. Related SoC-oriented efforts similarly report that careful workload partitioning and integration choices are necessary to obtain meaningful application-level speedups, particularly when standard-compliant encoding/decoding and signing-specific control flow (including rejection checks) remain in the software domain [23,24].

Overall, existing FPGA-based Dilithium research demonstrates strong kernel-level acceleration, yet it also exposes a practical gap that is central to edge computing: peak-throughput or high-volume designs do not automatically translate into lightweight, low-overhead deployments on compact SoC-FPGA platforms. NTT-focused engines [2,11] improve the arithmetic backbone, high-volume accelerators [5] clarify the importance of memory throughput and invocation overhead, and HW/SW co-designs [14] show that end-to-end gains hinge on a disciplined partitioning and a communication-efficient data path. These observations motivate the approach of this paper, which targets a lightweight ML-DSA signature engine for FPGA-based edge computing by jointly optimizing the NTT-centric datapath and the system integration costs that govern practical signing and verification latency under constrained resources.

3 Proposed FPGA-Based Dilithium Core Architecture

In this section, we introduce our FPGA-based CRYSTALS-Dilithium signature accelerator for edge computing. We design the architecture following a hardware acceleration paradigm, so that the proposed core can be integrated with either hard processors (e.g., ARM cores in SoC-FPGAs) or soft general-purpose processors to support practical signing and verification in embedded and edge applications. The design focuses on a lightweight integration model, where compute-intensive kernels such as NTT/INTT-based polynomial arithmetic and SHAKE-based hashing are mapped to the programmable logic, while control-oriented tasks and protocol orchestration are handled by software to preserve flexibility and standard compliance [17].

3.1 Overview Architecture

Figure 1 shows the general architecture of the proposed lightweight CRYSTALS-Dilithium post-quantum signature core designed for FPGA-based edge computing. The core communicates with a host processor through a communication infrastructure (e.g., AXI4 or shared memory) to transfer input data (entropy seeds, public keys, secret keys, messages, and parameters) and output results (generated keys, signatures, and verification decisions). A central “Internal BRAM” is utilized for storing and exchanging intermediate polynomials and data between all modules.

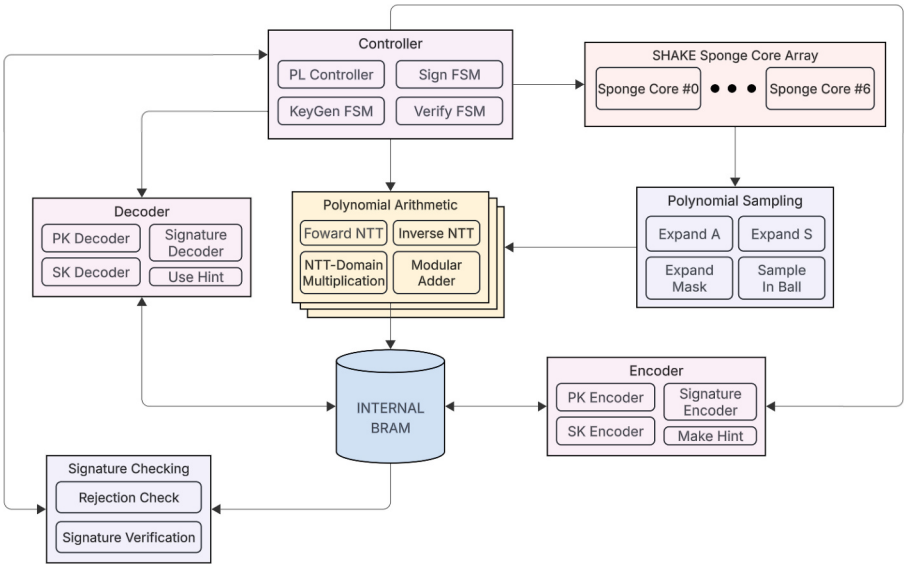


Fig. 1. The overview of the lightweight CRYSTALS-Dilithium Post-Quantum Cryptography Signature for FPGA-based Edge Computing.

The overall operation of the accelerator is managed by a global controller, which includes dedicated finite-state machines (FSMs) for the KeyGen, Sign, and Verify phases. Additional supporting modules, such as encoder/decoder units and signature check logic, handle the data formatting, compression, and validation steps required by the specification.

Among the processing components, two modules play a central role in accelerating the computationally intensive operations of the Dilithium algorithm: the polynomial arithmetic module and the polynomial sampling module.

The polynomial arithmetic module implements the core polynomial operations required by the scheme, including forward and inverse Number Theoretic Transform (NTT/INTT), modular addition and pointwise polynomial multiplication in the NTT domain.

The polynomial sampling module generates the pseudorandom polynomials required by the protocol. The implementations of the sampling procedures were defined in the Dilithium specification (FIPS 204), which derive structured randomness using extendable-output functions based on SHA-3. The detailed architectures of the polynomial arithmetic and polynomial sampling modules are discussed in the following subsections.

3.2 NTT Module

The proposed NTT core adopts a 2×2 radix-2 butterfly architecture, where two consecutive radix-2 butterfly units are arranged to process four input coefficients

in a pipelined manner. As illustrated in the architecture, the first stage consists of two parallel radix-2 butterflies that operate on pairs of inputs (x_1, x_2) and (x_3, x_4). Before entering the butterfly computation, the second operand of each pair is multiplied by the corresponding twiddle factor ζ^i. Each butterfly then performs modular addition and subtraction, generating two intermediate outputs according to Eq. 1.

$$\begin{aligned} t &= \zeta^i \cdot b \\ a' &= a + t \pmod{q}, \qquad b' = a - t \pmod{q}. \end{aligned} \tag{1}$$

The outputs of these two butterflies are then forwarded to the next stage, forming the 2×2 radix-2 structure. In the second stage, another pair of radix-2 butterflies receives the intermediate results from the first stage and performs additional modular operations using updated twiddle factors ζ^{2i+1} and ζ^{2i+2}. This cascading structure allows for four coefficients to be cooperatively processed, enables high-throughput processing, and reduces the complexity of the control. Figure 2 illustrates the NTT processing.

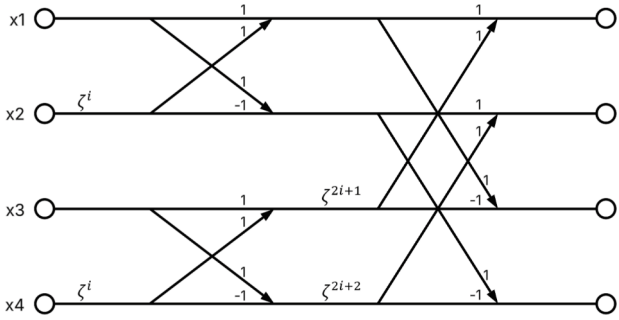


Fig. 2. Proposed 2×2 radix-2 butterfly architecture for the NTT core.

Based on the NTT computation principle, the architecture of the proposed NTT core is designed using a pipelined processing model to improve performance, as presented in Fig. 3. The computation is organized into several pipeline stages including modular multiplication, modular reduction, and butterfly operations. Pipeline registers are inserted between functional units to maintain continuous data flow and increase the operating frequency of the system.

The input coefficients are first received through the `Data in` interface and stored in pipeline registers. In the first computation stage, the input data is multiplied by the corresponding twiddle factor ζ provided by the input coefficient. The multiplication results are then passed through a modular reduction block to compute the modulo result q, ensuring that the values remain within the finite field used by the NTT. After modular reduction, intermediate results are forwarded to the butterfly computation stage, which performs modular addition and subtraction operations. The architecture then repeats a similar sequence

of operations in the subsequent stage, including multiplication, modular reduction, and butterfly computation, corresponding to the cascaded radix-2 butterfly structure.

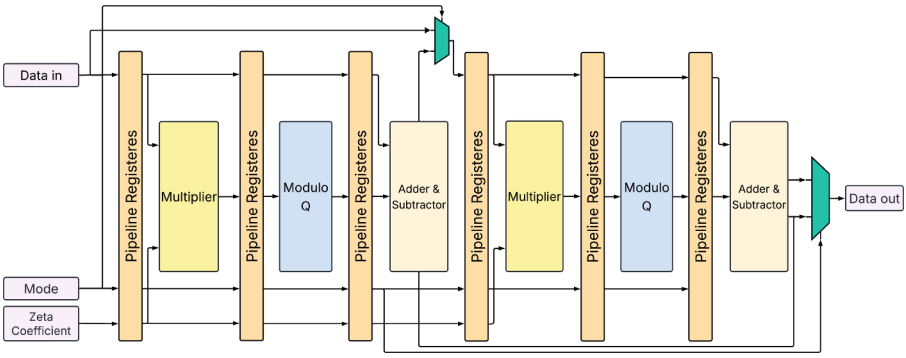


Fig. 3. The pipeline architecture for 2×2 radix-2 butterfly.

Multiplexers are placed at specific points in the data path to support different operation modes and to select appropriate intermediate results during the computation process. The pipeline registers between stages ensure that each functional block operates concurrently, allowing multiple data samples to be processed simultaneously in different stages of the pipeline. Finally, the processed coefficients are forwarded through the output multiplexer and written to the `Data out` interface.

Figure 4 shows the architecture of the polynomial processing block. The Computation Controller coordinates the overall operation, including input handling, data flow control, and output generation. An Address Generation Unit produces the RAM write addresses and twiddle-factor addresses needed for each processing stage. These addresses are used to access Zeta ROM, which stores the twiddle coefficients required by the NTT computation. The selected coefficients are then fed into the RADIX 2×2 block, where four butterfly units perform the core polynomial arithmetic. For inverse NTT processing, the intermediate results are forwarded to the Post-Inv NTT block, which applies the scaling constant, then passes the result to the Parallel-in Serial-out block for data formatting and serialization before being written back or delivered to the output interface. Through this organization, the architecture defines a clear data-path for polynomial computation, where the controller coordinates memory accesses, coefficient retrieved from the twiddle-factor ROMs, and the execution of arithmetic operations in the butterfly units.

3.3 Polynomial Sampling Module

The polynomial sampling module is responsible for generating the polynomials required by the Dilithium algorithms. As illustrated in Fig. 5, the module

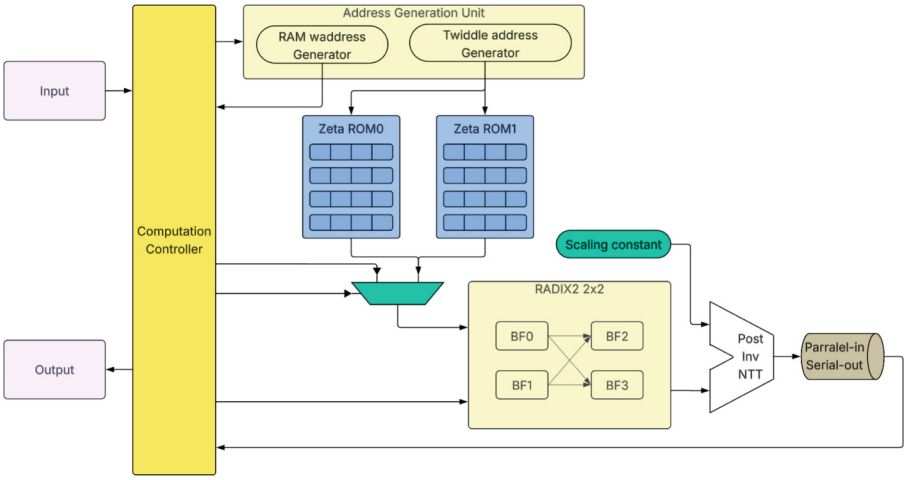


Fig. 4. The architecture of the whole operation module.

consists of three main components: the sampling controller, the SHAKE sponge engine, and the coefficient sampling unit. These components collectively implement the sampling procedures used in the ExpandA, ExpandS, ExpandMask, and SampleInBall operations.

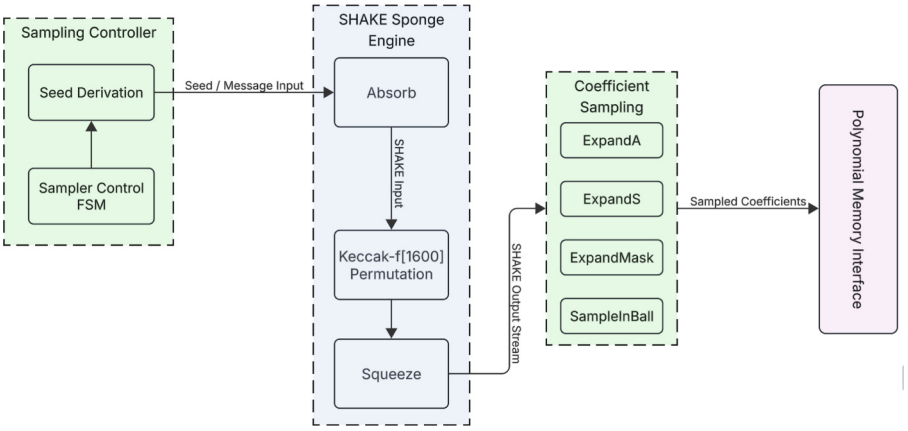


Fig. 5. The architecture for the polynomial sampling module.

The Sampling Controller operates the overall sampling process. It includes a Seed Derivation unit and a Sampler Control finite-state machine FSM. The Seed Derivation unit prepares the deterministic seeds required by the SHAKE engine based on the input parameters, such as the public seed, secret seed, message digest, and nonce values specified in the Dilithium algorithms. These seeds

are forwarded to the SHAKE engine during the absorption phase. The sampler control FSM then selects and activates the appropriate sampling procedure depending on the current algorithm phase (e.g., matrix expansion during key generation or mask sampling during signing).

The pseudorandom streams used for coefficient generation are produced by a SHAKE sponge engine based on the Keccak-f[1600] permutation. The engine follows the standard sponge construction with absorb, permutation, and squeeze phases. After absorbing the input, the permutation is applied to update the internal state, and the squeeze phase generates a pseudorandom output bitstream. This stream serves as the entropy source for the coefficient sampling unit. In the proposed implementation, the engine can be instantiated as SHAKE128 or SHAKE256.

The coefficient sampling unit converts the SHAKE output stream into polynomial coefficients according to the Dilithium specification. The ExpandA operation deterministically expands a seed into public matrix A. The ExpandS operation generates secret vector polynomials with small coefficients, while ExpandMask produces masking polynomials used during the signing process. The SampleInBall procedure generates a sparse challenge polynomial with a fixed Hamming weight. The sampled coefficients are written to the polynomial memory interface, allowing subsequent modules such as the polynomial arithmetic unit to access the generated data.

4 FPGA Edge Computing Prototype

This section presents the prototype FPGA-based edge computing platform developed for integrating a CRYSTALS-Dilithium post-quantum cryptographic accelerator on the Xilinx Kria KV260 platform. The prototype is implemented on the Xilinx Zynq UltraScale+ MPSoC, where the system is divided into two main parts: the Processing System (PS) and the Programmable Logic (PL). The accelerator is described in Verilog HDL and synthesized using Xilinx Vivado, while the embedded ARM Cortex-A53 processor in the PS serves as the host processor to control and interact with the hardware accelerator rewrite this paragraph with the same length but change the expression way.

Figure 6 illustrates the overall architecture of the proposed system. In the Processing System, the ARM Cortex-A53 processor running PYNQ is responsible for parameter initialization, user interaction, data pre-processing, memory read/write operations, and validation of hardware outputs against a golden software model. The processor communicates with the accelerator in the PL through the M_AXI_HPM0_FPD AXI master interface, which is connected to AXI SmartConnect. Inside the PL, the AXI SmartConnect links both the Dilithium accelerator core and the AXI BRAM Controller, while the dual-port RAM acts as a shared on-chip memory buffer between the processor and the accelerator.

The Dilithium accelerator core consists of several functional modules. An AXI slave interface provides the control and status interface between the host processor and the accelerator. A Control FSM coordinates the execution flow

of the core and dispatches operations to the cryptographic engines. The core includes dedicated KeyGen, Sign, and Verify engines, which implement the main Dilithium operations. These engines rely on an Encoder/Decoder block for packing and unpacking keys, signatures, and intermediate data, a Polynomial Arithmetic Unit for polynomial and modular arithmetic operations, and a Polynomial Sampling block for generating the random and structured polynomials required by the Dilithium algorithm.

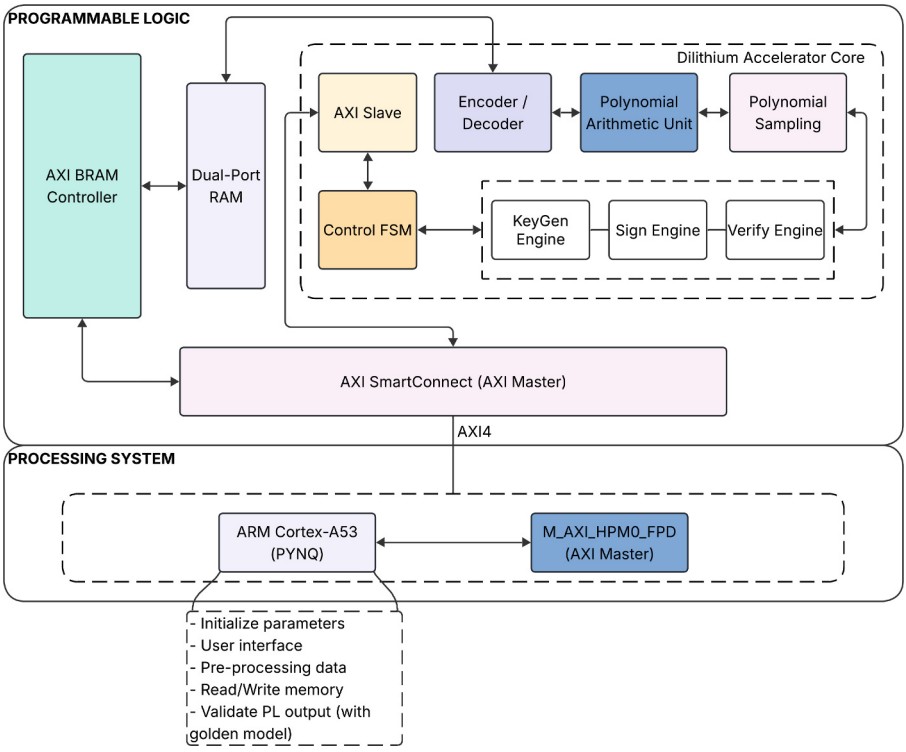


Fig. 6. The prototype system with Xilinx Kria KV260 board.

During operation, the host processor initializes the required parameters and input data, stores them in the shared BRAM, and triggers the accelerator through the AXI slave control interface. The Control FSM then activates the corresponding KeyGen, Sign, or Verify engine, while the arithmetic, encoding/decoding, and sampling modules support the requested computation. After execution is completed, the generated outputs are stored back into memory and read by the host processor for post-processing or correctness verification.

5 Experiments

This section presents experiments to evaluate the effectiveness of the CRYSTALS-Dilithium accelerator core and the FPGA-based prototype systems. The experiment results are then examined, comparing the system's performance with the accelerator against using only software on the ARM processor. In addition, power consumption will be evaluated, analyzed, and compared.

5.1 Synthesis Results

The CRYSTALS-Dilithium accelerator core, developed by Verilog-HDL, and the entire system are synthesized and configured on the Kria board by Xilinx Vivado 2024.2. The results of the synthesis reveal that the system can operate at 133 MHz and requires 3.71 W of total power consumption. Table 1 presents the usage of hardware resources for the system.

Table 1. Hardware resources usage for the proposed system

Resources	Amount	Percentage	Available
Look-up tables (LUTs)	97,215	83%	117,120
Flip-flop (FFs)	64,913	27.71%	234,240
DSP	48	3.85%	1,248
Block RAMs (BRAMs)	89	61.8%	144
$F_{max} = 133$ MHz		$P = 3.71$ W	

As presented in Table 1, the proposed system uses 97,215 LUTs (83% of the available resources) and 64,913 FFs (27.71%), indicating that the design is primarily logic-intensive. The relatively high utilization of LUT results mainly from the implementation of polynomial arithmetic units and multiple SHAKE sponge cores used for polynomial sampling. In contrast, the usage of DSP blocks and BRAMs remains moderate. This shows that the proposed architecture efficiently leverages FPGA logic resources while maintaining reasonable consumption of dedicated hardware blocks.

5.2 Simulation Results

Figure 7 shows the timing waveform of the butterfly core 1× 1 operating in *FORWARD_NTT* mode. When `validi` is asserted, the input coefficients and twiddle factors are processed through three pipeline stages: twiddle multiplication, modular reduction, and butterfly addition/subtraction. After the pipeline latency, the output coefficient pairs are produced sequentially, and `valido` indicates valid output data.

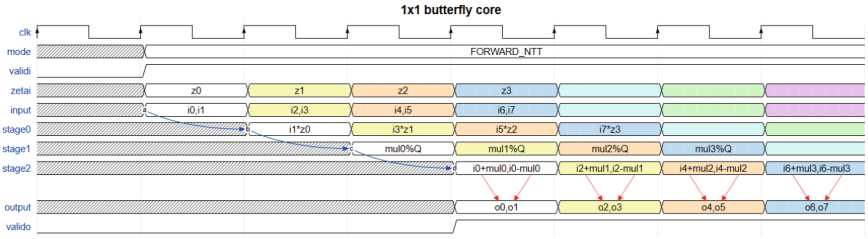


Fig. 7. The waveform for simulation 1×1 butterfly core.

Following the timing behavior of the 1×1 butterfly core shown previously, Fig. 8 presents the waveform of the proposed 2×2 radix-2 butterfly core operating in *FORWARD_NTT* mode. In this architecture, the computation is performed in two butterfly stages. When `validi` is asserted, the input coefficients and the corresponding twiddle factors are first processed by two parallel butterflies (BF0 and BF1), producing intermediate results t_0, t_1, t_2, t_3. Before being transferred to the second stage, the intermediate outputs are rearranged by swapping the positions of t_2 and t_3 to match the required input ordering of the following butterfly units. The reordered values are then processed by butterflies BF2 and BF3 using the corresponding twiddle factors. Due to the pipelined structure, the intermediate results appear first in the output of the first stage, followed by the final output coefficients generated by the second stage. The signals `valido_bf01` and `valido_bf23` indicate the availability of valid intermediate and final results, respectively.

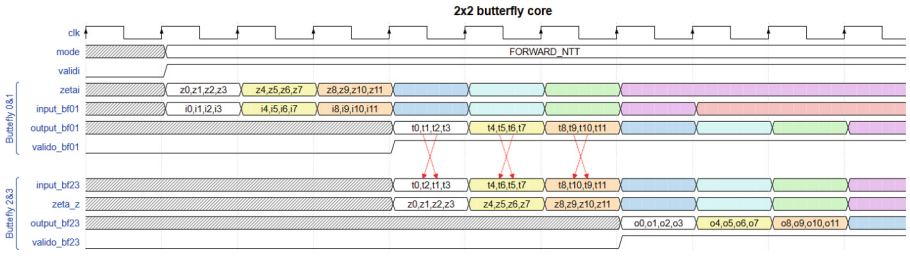


Fig. 8. The waveform for simulation 2×2 butterfly core.

For polynomial sampling, Fig. 9 illustrates the timing waveform of the module `ExpandA`, chosen as a representative example among the four polynomial generation procedures. The SHAKE engine continuously produces pseudorandom outputs (shake_valid, shake_data), which are processed to generate candidate coefficients. Due to the rejection sampling procedure, invalid candidates are discarded, causing the coefficient index (coeff_idx) and the packing counter (coeff_cnt) to temporarily stall. Valid coefficients are accumulated and four coefficients are packed into a single memory word before being written into polynomial

memory when write_en is asserted. After all coefficients of the current polynomial are generated, the polynomial index (poly_idx) increases and the sampling process continues for the next polynomial.

Fig. 9. The waveform for simulation ExpandA module.

5.3 Performance Evaluation and Comparison

We evaluate the performance of the proposed processor using message sizes ranging from 32 bytes to 20 kB for the signing and verification procedures. The evaluation compares two implementations: (1) A hardwired ARM core operating at 1.33 GHz running the reference Dilithium implementation written in C (Software); and (2) the proposed accelerated system, which integrates a dedicated CRYSTALS-Dilithium accelerator core (FPGA).

Table 2 summarizes the execution time of the proposed accelerator system compared to a software implementation running on an ARM processor operating at 1.33 GHz. The proposed design achieves speedups of up to 3.66×, 4.56× and 4.18× for the key generation, signing and verification phases, respectively.

The main performance improvement originates from the hardware acceleration of polynomial arithmetic operations, including number theoretic transform (NTT), inverse NTT, and polynomial multiplications. These operations dominate the computational workload in Dilithium and benefit significantly from the parallel datapath implemented in the accelerator. For evaluating polynomial arithmetic modules, a standard 1×1 polynomial matrix is used as the baseline input to measure the latency of NTT, INTT and polynomial multiplication operations. Despite the substantial acceleration of polynomial arithmetic, the overall speedups of the main phases remain moderate. This is primarily because several components of the Dilithium algorithm, such as the rejection sampling procedures and the SHAKE hash function, are less amenable to hardware acceleration while accounting for a considerable fraction of the total execution time.

Table 3 shows that the proposed NTT architecture achieves competitive latency compared to recent FPGA-based implementations. In particular, this work obtains 1.93 μs for both NTT and INTT, which is faster than most previous designs on Zynq-7000, Artix-7, and ZCU-series platforms in the literature and is close to the best reported results. For polynomial-wise multiplication, the proposed design achieves 0.48 µs, outperforming the compared works reporting

Table 2. Execution time (μs) of each scheme

Scheme	FPGA		Software
	Exec. time	Speed-up	
ExpandA	238.71	2.79×	665.26
ExpandS	47.96	2.68×	128.66
ExpandMask	27.98	3.45×	96.63
SampleInBall	1.48	3.29×	4.87
Foward NTT	1.93	11.79×	22.76
Inverse NTT	1.93	13.56×	26.18
Polynomial Multiplication	0.48	11.10×	5.33
Key Generation	398.61	3.66×	1461.53
Sign	1356.41	4.56×	6181.28
Verify	392.83	4.18×	1642.41

this metric. These results indicate that the proposed architecture provides an effective balance between lightweight implementation and high computational efficiency for edge-oriented Dilithium acceleration.

Table 3. Performance comparison of NTT design with other implementations

Work	Year	Platform	Latency (μs)		
			NTT	INTT	PWM
Gaoyu Mao [14]	2023	Zynq-7000	8.17	8.17	1.51
Tengfei Wang [24]	2024	PolarFire	1.96	1.96	0.54
Stefano Di Matteo [7]	2024	ZCU104	3.91	3.91	-
Suraj Mandal [13]	2024	Zynq 7000	3.67	3.67	0.92
Suraj Mandal [13]	2024	Artix-7	3.75	3.75	0.94
Suraj Mandal [13]	2024	ZCU102	2.48	2.48	0.62
Rishabh Shrivastava [22]	2025	ZCU104	1.56	1.56	-
This Work	**2026**	**Kria KV260**	**1.93**	**1.93**	**0.48**

6 Conclusion

This paper presented a lightweight FPGA-based hardware accelerator for the CRYSTALS-Dilithium post-quantum digital signature scheme targeting edge computing environments. The proposed architecture integrates a dedicated Dilithium accelerator core within the programmable logic of the Xilinx Kria

KV260 platform, while the ARM Cortex-A53 processor in the processing system acts as the host controller. Using hardware parallelism and optimized polynomial arithmetic units, the proposed design accelerates key cryptographic operations, including key generation, signing, and verification. Experimental results demonstrate that the hardware accelerator significantly improves execution efficiency compared with software-based implementations while maintaining a compact resource footprint suitable for edge devices. These results confirm the feasibility of deploying post-quantum digital signature mechanisms on FPGA-based edge platforms to enhance security in resource-constrained environments.

Acknowledgement. This research is funded by Vietnam National University - Ho Chi Minh City (VNU-HCM) under grant number: A2025-20-01.

References

1. Baseri, Y., Chouhan, V., Ghorbani, A., Chow, A.: Evaluation framework for quantum security risk assessment: a comprehensive strategy for quantum-safe transition. Comput. Secur. **150**, 104272 (2025)
2. Beckwith, L., Nguyen, D.T., Gaj, K.: High-performance hardware implementation of crystals-dilithium. In: 2021 International Conference on Field-Programmable Technology (ICFPT), pp. 1–10. IEEE (2021)
3. Boutin, C.: NIST Releases First 3 Finalized Post-Quantum Encryption Standards (2024). https://www.nist.gov/news-events/news/2024/08/nist-releases-first-3-finalized-post-quantum-encryption-standards
4. Brakerski, Z., Langlois, A., Peikert, C., Regev, O., Stehlé, D.: Classical hardness of learning with errors. In: Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing, pp. 575–584 (2013)
5. Carril, X., Kardaris, C., Ribes-GonzáLez, J., Farràs, O., Hernandez, C., Kostalabros, V., González-Jiménez, J.U., Moreto, M.: Hardware acceleration for high-volume operations of crystals-kyber and crystals-dilithium. ACM Trans. Reconfigurable Technol. Syst. **17**(3), 1–26 (2024)
6. Chen, L., Moody, D., Regenscheid, A., Robinson, A.: Digital signature standard (2023)
7. Di Matteo, S., Sarno, I., Saponara, S.: Cryphtor: a memory-unified ntt-based hardware accelerator for post-quantum crystals algorithms. IEEE Access **12**, 25501–25511 (2024)
8. Halak, B., et al.: A security assessment tool for quantum threat analysis (2024). arXiv:2407.13523 arXiv preprint
9. Hsu, R.-H., Lee, J., Quek, T.Q.S., Chen, J.-C.: Reconfigurable security: edge-computing-based framework for iot. IEEE Network **32**(5), 92–99 (2018)
10. Jin, X., Katsis, C., Sang, F., Sun, J., Kundu, A., Kompella, R.: Edge security: Challenges and issues. arXiv preprint arXiv:2206.07164 (2022)
11. Kundi, D.E.S., Mera, J.M.B., Strub, P.Y., Hutter, M.: High-performance ntt hardware accelerator to support ml-kem and ml-dsa. In: Proceedings of the 2024 Workshop on Attacks and Solutions in Hardware Security, pp. 100–105 (2024)
12. Langlois, A., Stehlé, D.: Worst-case to average-case reductions for module lattices. Des. Codes Crypt. **75**(3), 565–599 (2014). https://doi.org/10.1007/s10623-014-9938-4

13. Mandal, S., Roy, D.B.: Kid: a hardware design framework targeting unified ntt multiplication for crystals-kyber and crystals-dilithium on fpga. In: 2024 37th International Conference on VLSI Design and 2024 23rd International Conference on Embedded Systems (VLSID), pp. 455–460. IEEE (2024)
14. Mao, G., Chen, D., Li, G., Dai, W., Sanka, A.I., Koç, Ç.K., Cheung, R.C.: High-performance and configurable sw/hw co-design of post-quantum signature crystals-dilithium. ACM Trans. Reconfigurable Technol. Syst. **16**(3), 1–28 (2023)
15. Mascelli, J., Rodden, M.: "Harvest Now Decrypt Later": Examining Post-Quantum Cryptography and the Data Privacy Risks for Distributed Ledger Networks (2025)
16. Moody, D., Perlner, R., Regenscheid, A., Robinson, A., Cooper, D.: Transition to post-quantum cryptography standards. Tech. rep, National Institute of Standards and Technology (2024)
17. NIST: FIPS 204: Module-Lattice-Based Digital Signature Standard. Tech. rep., U.S. Department of Commerce (2024). https://doi.org/10.6028/NIST.FIPS.204, Federal Information Processing Standards Publication 204
18. Peikert, C.: A decade of lattice cryptography. Found. Trends Theoret. Comput. Sci. **10**(4), 283–424 (2016)
19. Regev, O.: On lattices, learning with errors, random linear codes, and cryptography. J. ACM (JACM) **56**(6), 1–40 (2009)
20. Sha, K., Yang, T.A., Wei, W., Davari, S.: A survey of edge computing-based designs for iot security. Digital Commun. Networks **6**(2), 195–202 (2020)
21. Shor, P.W.: Algorithms for quantum computation: discrete logarithms and factoring. In: Proceedings 35th Annual Symposium on Foundations of Computer Science, pp. 124–134. IEEE (1994)
22. Shrivastava, R., Ratnala, C.P., Puli, D.M., Banerjee, U.: A unified hardware accelerator for fast fourier transform and number theoretic transform. In: ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1–5. IEEE (2025)
23. Wang, T., Zhang, C., Cao, P., Gu, D.: Efficient implementation of dilithium signature scheme on fpga soc platform. IEEE Trans. Very Large Scale Integr. VLSI Syst. **30**(9), 1158–1171 (2022)
24. Wang, T., Zhang, C., Zhang, X., Gu, D., Cao, P.: Optimized hardware-software co-design for kyber and dilithium on risc-v soc fpga. IACR Trans. Cryptographic Hardware Embedded Syst. **2024**(3), 99–135 (2024)
25. Xue, T., et al.: Edge Computing for IoT: Novel Insights from a Comparative Analysis of Access Control Models. arXiv preprint arXiv:2405.07685 (2024)
26. Zheng, J., Zhu, H., Song, Z., Wang, Z., Zhao, Y.: Optimized vectorization implementation of crystals-dilithium. IEEE Trans. Dependable Secure Comput. (2025)

A Binning-Based Approach for Intrusion Detection and Attack Classification in Internet of Medical Things Networks

Lerina Aversano, Stefano Galantucci(✉), Michele Mastroianni, and Andrea Porcelli

Department of Agricultural Sciences, Food, Natural Resources and Engineering (DAFNE), University of Foggia, Foggia 71122, Italy
{lerina.aversano,stefano.galantucci,michele.mastroianni, andrea.porcelli}@unifg.it

Abstract. Internet of Medical Things (IoMT) is redefining modern healthcare by enabling the continuous collection, transmission, and analysis of patient data through interconnected medical devices, wearables, and hospital systems. These advantages, however, extend and complicate the e-health attack surface, the effects of which are devastating as they impact human health and privacy. Reports indicate a rise in IoT-based cyberattacks targeting medical infrastructures and sensitive data regarding patients' health; consequently, Intrusion Detection and Prevention Systems (IDS/IPS) have become essential components of secure IoMT architectures. IoMT networks generally have different characteristics from traditional networks and require ad hoc solutions. Most IDS/IPS models in the literature rely on deep learning architectures that, while achieving high accuracy, are unsuitable for real-time operation in resource-constrained IoMT environments. The resulting gap between recognition performance and solution deployability leaves IoMT networks lacking real-world applicable approaches. To address this gap, this paper introduces a lightweight hybrid IDS/IPS framework that combines decision tree models with a compact neural network. The neural network is enhanced through a supervised binning mechanism designed to handle non-continuous and heterogeneous IoMT data effectively. Experiments on three widely adopted IoMT datasets (CICIoMT 2024, IoMT Traffic-Data, and WUSTL EHMS 2020) show that the proposed model achieves accuracy and F1-weighted scores exceeding 0.99, while processing 30,000–100,000 samples per second. Compared to state-of-the-art deep models, our approach maintains comparable or superior detection capability with significantly lower computational cost. The model proved to be particularly robust and effective in attack detection (binary classification) and attack classification (multiclass classification) tasks.

Keywords: Internet of Medical Things (IoMT) · Internet of Things (IoT) · Intrusion Detection and Prevention Systems (IDS/IPS) · E-health · Healthcare · Binning · Decision Tree · Neural Network

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 351–368, 2027.
https://doi.org/10.1007/978-3-032-30530-5_22

1 Introduction

The integration of interconnected devices into the healthcare domain, known as the Internet of Medical Things (IoMT) [4,26,54], is revolutionizing patient care by enabling continuous real-time monitoring and remote management of clinical data. These systems connect medical devices, wearables, gateways, and hospital information systems across wards, ambulances, and homes, so that data can flow where it is needed for timely decisions and safe treatment [10,39]. Modern IoMT architectures offer unprecedented advancements in personalized healthcare solutions, the continuous transmission of highly sensitive physiological data introduces severe privacy vulnerabilities. This rapid expansion, introduces critical cybersecurity vulnerabilities [13,15,44]. According to data obtained from SonicWall Capture Labs in 2022, over 112 million IoT cyber attacks have been reported worldwide [1]. The sensitive nature of health data and the direct impact on patient safety and personal information make IoMT systems a prime target [30,48] for attacks that can disrupt or degrade essential services. Attacks on IoMT systems can lead to catastrophic consequences. Breaches of confidentiality can result in the theft of health data, breaches of availability can cause the failure of devices that help to ensure patient health, and breaches of integrity can alter the functioning of devices, causing them to operate improperly. In the best-case scenario, this results in a serious privacy violation, while in the worst-case scenario, it can lead to the death of a device-using patient. In this context, the development of effective Intrusion Detection/Prevention systems (IDS/IPS) capable of identifying anomalies in real time is not merely a technical task but a requirement for keeping connected healthcare reliable and secure [22,47,59]. Over the years, a wide range of solutions for IDS/IPS systems has been developed [20,21,36,42,45,49,55,56], but these are aimed at generic or corporate networks. IoMT networks and, more generally, IoT networks exhibit different characteristics [14,17,58], different protocols [12,19], and different traffic flows and volumes. This implies that this research problem requires different approaches: from this perspective, research has explored many learning-based IoMT IDS/IPS solutions. Many works [11,16,18,19,29,38] propose deep learning models to capture complex, non-linear interactions in traffic and device behavior. For example, convolutional networks are used to learn temporal patterns [32,43], recurrent models [53] to capture longer dependencies, and newer architectures that even transform tabular data into image-like structures to be processed by vision models [3,7,53]. Other authors propose ensemble approaches [28,34,37] that combine different deep models to push accuracy as high as possible. However, these advances often come with trade-offs. Deep models achieve good performance but are slow to train and heavy to execute. Conversely, simple deep models on tabular data, although lightweight, often struggle to reach high scores [11,19]. The constraints of time and resources required are indeed fundamental: on the one hand, time can make a difference to the patient's life; on the other hand, IoMT solutions are not generally equipped with large computational resources, so they cannot easily handle heavy approaches. Authors in [6,28] emphasise the need for fast, lightweight solutions that guarantee real-time monitoring and

detection. IoMT supports "real-time monitoring" of patients [51]; consequently, IDS/IPS must also operate in real-time to be effective. In short, IoMT intrusion detection should balance effectiveness with practicality. This alignment between algorithm and context is key to building secure, trustworthy, and sustainable IoMT systems. To address these needs, this work proposes a lightweight but highly effective IDS/IPS solution for IoMT networks. The proposed approach consists of a combination of decision trees and a simple neural network. These families were chosen because trees have proven to be highly effective on IoMT data [11,19], particularly on tabular telemetry data [28,40], and are easy to inspect, while small neural networks can model non-linear patterns with modest computing requirements [38]. The goal is not only to maximize accuracy on a benchmark, but to deliver dependable detection that fits IoMT workflows and meets resource limits. The proposed approach has been tested on three datasets widely used in the IoMT literature [9,11,19], obtaining results that, in most cases, outperform those reported in the current literature. Furthermore, since the literature mainly proposes heavyweight approaches, our model guarantees real applicability even on small medical devices, providing a real increase in the level of security in the e-health field.

The key contributions of this work are:

- the proposal of an IDS/IPS methodology for detecting and classifying attacks in IoMT networks, based on trees and neural networks;
- the testing of this methodology on three widely-used IoMT datasets: CICIoMT2024 [19], IOMT-TrafficData [11], and WUSTL-EHMS 2020 [9];
- comparison of the results obtained with the current literature, with respect to standard metrics and model weighting.

This work is organized as follows: the proposed approach is described in Sect. 3, the related work are presented in Sect. 2, the experiments and related results are discussed in Sect. 4, and the conclusions are in Sect. 5.

2 Related Work

The growing diffusion of IoMT requires robust and efficient IDS, capable of mitigating ever-increasing threats, safeguarding sensitive patient data, and ensuring system integrity from vulnerabilities or attacks [13,15,44]. Research directions focus on enhancing accuracy and generalization capabilities, often overlooking important metrics such as inference time and training cost. The balance between metrics and computational efficiency, which distinguishes between fast (low-latency) and slow (high-latency) security solutions, characterizes this research field. Indeed, the main solutions can be split into two groups: "heavy" solutions that aim for maximum performance and "lightweight" solutions that seek a compromise. The primary area for the proposed work concerns the development of robust and fast solutions, essential for real-time monitoring and for IoMT devices with limited resources [32,34].

The subsequent paragraphs discuss machine learning models that are "lightweight" and commonly use aggressive feature engineering. This category includes Recursive Feature Elimination (RFE), used by [40] with 0.979 in test accuracy, 0.965 precision, and 0.863 recall with a DT. It also includes mixed approaches like RFE/MLP [38], which aim to minimize feature dimensionality with RFE and, at the same time, optimize the MLP, achieving 0.962 in every metric, consequently reducing computational overhead during training and inference. These two works use the dataset [9].

Similarly, the work of [33] highlights the effectiveness of feature selection based on Information Gain (IG) on the CICIoMT2024 dataset. This achieves with a MultiD-CNN the highest score in 5-fold cross-validation compared to many models, with a mean accuracy of 0.981, a mean precision of 0.977, a mean recall of 0.982, and a mean F1-score of 0.979. IG simplifies the dataset and improves classifier performance, but the MultiD-CNN model used exhibits high latency compared to other approaches.

The use of custom feature selection techniques based on mutual information, such as LRGU-MIFS [5], further emphasizes the push towards streamlining data complexity, reaching 0.949 in accuracy on the WUSTL-EHMS 2020 dataset [9]. Authors in [12] studied the behavior of these particular trees (Hoeffding DT) and other models on a dataset synthetically generated using IoT-Flock [25]. Hosain et al. [32] used XGBoost and XAI on WUSTL-EHMS 2020 [6], reaching the high accuracy score of 0.992. The approach proposed by Gupta et al. [28], which incorporates data augmentation on a random forest, has demonstrated exceptionally low latency on WUSTL-EHMS 2020 [9] with 0.942 accuracy, proving highly suitable for rapid detection scenarios. The Nilsimsa fingerprinting approach [31] provides a simplified modeling method for network traffic by bypassing the extensive feature extraction and selection phases commonly required by ML methods and using a hashing function, thereby increasing efficiency and achieving 0.982 accuracy on the CICIoMT2024 dataset. The goal of developing systems that are efficient in terms of resources and computational complexity, often incorporating lightweight ML classifiers such as Logistic Regression (LR) and K-Nearest Neighbors (KNN) into ensemble methods, remains fundamental for their implementability [9,34].

Conversely, in the literature, there are numerous Deep Learning (DL) architectures that often represent slower and more resource-intensive solutions, designed for maximum accuracy and to handle complex dependencies. Deep Learning (DL) models, while demonstrating superior performance in multi-class classification tasks (e.g., distinguishing 19 attack subtypes on the CICIoMT2024 dataset), frequently involve high prediction times or significant training overhead [16,29]. Examples include Convolutional Neural Network (CNN) models optimized for temporal analysis [43], achieving on CICIoMT2024 [19] 0.990 accuracy, 0.980 precision, 0.990 recall, 0.980 F1-score in multi-class classification, and 0.990 over all metrics in binary, demonstrating high performance although the convolutional layers are heavy. Deep Neural Networks (DNN) [29], and methods employing optimization algorithms such as Particle Swarm Optimization (PSO)

used in PSO-DNN models [18] on WUSTL-EHMS 2020, achieving 0.96 accuracy with a DNN, are recognized as computationally expensive [38,40]. Additionally, ensemble models are used [27]. Although the well-known cost in terms of computation in the IoT/IoMT environment, where efficiency is a must, evaluating the ten learning methods utilizing the TON_IoT [8] attack dataset, obtaining Matthews correlation coefficient (MCC) scores of 0.9971. The stacking-ensemble model emerged as the top performer. Highly advanced sequence modeling techniques, such as the use of Vision Transformers (ViT) in threat detection [7] reaching 0.995 across all metrics on the CICIoMT2024 dataset can be interesting from a theoretic point of view but unfeasible in real-world scenarios, or hybrid Recurrent Neural Networks (RNNs) like the BiGRU-BiLSTM proposed by [53] (0.999 on IoMT-TrafficData), CNN-LSTM hybridization proposed by [23] on [52], or the LSTM-based L2D2 (0.98 across all metrics on the CICIoMT2024 [19] dataset) model aiming for an optimal balance between accuracy and computational demand [3], yield high performance but increase architectural complexity and training time [38,46]. Features complexity reduction approaches drop irrelevant and redundant features from the dataset, enhancing model performance. However, efforts are being made to mitigate the latency of these models post-training through techniques such as model quantization, demonstrating low inference times even for complex hybrid DL models [53]. A critical research direction is the integration of these sophisticated models into distributed and privacy-preserving security architectures. This implies moving beyond centralized analysis towards systems that incorporate concepts such as Federated Learning (FL) and Blockchain [15,53]. The work of [37] proposes a multi-view learning approach within an FL setup to address scalability challenges and non-IID (non-independent and identically distributed) data inherent in real-world healthcare environments with an impressive accuracy score of 0.994 on the CICIoMT2024 [19] dataset. Migrating IDS models to real-world IoMT requires robust validation [32,34] and strict avoidance of data leakage [16]. However, current literature faces a persistent dichotomy: lightweight ML struggles with complex attacks, while deep learning (DL) imposes prohibitive computational overheads for edge deployment. Moreover, both frequently mishandle the heterogeneous, non-continuous nature of IoMT tabular data. Although feature discretization (*binning*) [24,41] is occasionally used in network security to simplify features [57], traditional unsupervised methods arbitrarily split data, discarding critical predictive signals. To bridge this gap, this paper introduces a hybrid solution leveraging a *supervised*, tree-based binning mechanism. By explicitly optimizing neural network performance on segmented data, our approach achieves DL-level accuracy while maintaining the minimal resource footprint vital for real-time medical device security.

3 Proposed Approach

The proposed approach is based on the hypothesis that the lower effectiveness [11,19] of simple neural networks (such as Multi-Layer Perceptrons, MLPs)

compared to tree-based algorithms in IoMT attack classification contexts is attributable to the intrinsic nature of certain features and to neural networks' difficulty in modeling non-continuous or piecewise decision functions. Recent studies [31,32,40] show that, especially with tabular data where features have heterogeneous semantics, tree-based algorithms are often more effective. The main idea behind the proposed approach is that this is not due to an intrinsic weakness of the neural model, but rather its inadequacy in natively handling categorical or numerical features that exhibit a "segmented" behavior—like IoMT data—without proper data preparation. The following examples help understand this issue:

- **Network Ports:** A feature representing a packet's destination port is categorical, as there is no numerical or ordinal relationship between port 80 (HTTP) and port 81; they are simply distinct identifiers. A decision tree can easily create a splitting rule like `port == 80` to isolate a specific behavior. A neural network, on the other hand, would struggle immensely if it received this feature as a single numerical value. It would have to learn, through linear combinations and non-linear activation functions (like ReLU), a complex function that isolates specific points on a numerical axis—an intrinsically difficult and inefficient task. The standard solution, one-hot encoding [50], can become impractical if the number of ports analyzed is very high.
- **Packet Length:** Although this is a continuous numerical variable, in certain contexts (like the Internet of Medical Things, IoMT), its behavior is not linear. Legitimate devices might communicate using packets within a very specific size range (e.g., 32–64 bytes). An attack, however, might use very small (e.g., < 10 bytes) or very large (e.g., > 1500 bytes) packets. This creates "zones of interest" that do not follow a linear progression. A decision tree can easily capture these intervals with rules like `32 < length < 64`. A neural network, in contrast, would have to model a function that behaves one way within an interval and another way outside of it, another task that requires the complex approximation of a step function.

In both cases, decision trees would achieve very high performance because their nature is to partition the feature space using thresholds and equalities, whereas neural networks are optimized for learning complex, continuous transformations of data. A possible solution proposed in this work is based on binning [35,41], which is a feature engineering process that converts continuous variables (x_j) into discrete representations (x'_j), also called feature discretization [24,57]. Binning can help in this particular context by improving robustness to outliers, decreasing overfitting, and handling important characteristics of features. This study employs a supervised, tree-based approach to create bins that are maximally informative about the target variable y.

The proposed methodology (illustrated in Fig. 1) is composed of the following phases:

1. **Ranking and selection of Top Features**: a preliminary Decision Tree T_{prelim} is trained on the full training set $D_{train} = (X_{train}, y_{train})$ to rank all

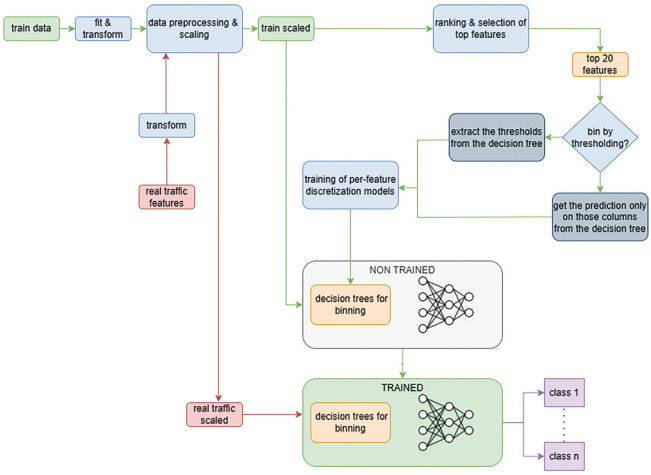


Fig. 1. Full operation of the proposed methodology.

features by their importance scores, I_j. Then, the top k features are selected to form the subset $X_{top\text{-}k}$;

2. **Training of Pre-Feature Discretization model**: for each feature $x_j \in X_{top\text{-}k}$, a dedicated, single-feature Decision Tree $T_j(\alpha)$ is trained. The model is trained using only the feature vector $\{x_j\}$ and the target vector y_{train};
3. **Feature Augmentation via Binning**: the trained trees $T_j(\alpha)$ are used to generate new binned features and augment the original datasets. This transformation is performed using two distinct techniques.
   - **Method A (Binning by Prediction)**: The new feature value is the tree's predicted class.

$$x'_{j,i} = T_j(\alpha).\text{predict}(x_{j,i})$$

   - **Method B (Binning by Thresholding)**: The new feature value is the index of the tree leaf where the sample lands.

$$x'_{j,i} = T_j(\alpha).\text{apply}(x_{j,i})$$

4. **Dataset preparation**: the binned feature set (e.g., $X_{\text{bin_pred, train}}$) are appended as new columns to the original feature matrix (e.g., X_{train}), thereby creating the final augmented feature matrix (e.g., $X'_{\text{pred, train}}$).
   - **For the Binning by Prediction method**:
     - Augmented **training set**: $X'_{\text{pred, train}} = [X_{train} \| X_{\text{bin_pred, train}}]$
     - Augmented **test set**: $X'_{\text{pred, test}} = [X_{test} \| X_{\text{bin_pred, test}}]$
   - **For the Binning by Thresholding method**:
     - Augmented **training set**: $X'_{\text{thresh, train}} = [X_{train} \| X_{\text{bin_thresh, train}}]$
     - Augmented **test set**: $X'_{\text{thresh, test}} = [X_{test} \| X_{\text{bin_thresh, test}}]$

where $\|$ denotes the horizontal concatenation of the feature matrices.

The binning in the third step can be interpreted from two perspectives:

1. **Threshold-Based Binning** [2]: From this viewpoint, each data point is allocated to a specific segment (or "bin") based on the path it takes through the tree's decision nodes. The internal nodes act as thresholds (for example: $X > 75$, $X > 100$, $X > 50$) that define the boundaries of these segments.
2. **Prediction-Based Binning:** This perspective focuses on the outcome. All data points that fall into the same segment are subsequently assigned the identical class label "like an ensemble" (e.g., 'malware' or 'goodware') defined by the terminal leaf node corresponding to that segment.

4 Experiments

This section details the comprehensive experimental design for this study. The hardware setup is provided in Sect. 4.1. The datasets employed are discussed in Sect. 4.2. The partitioning and validation strategy are provided in Sect. 4.3. The hyperparameter configurations are indicated in Sect. 4.4. The results of the experiments are reported and discussed in Sect. 4.5.

4.1 Hardware Setup

The experiments were conducted on a workstation equipped with an Intel Core i7-8700K processor, 32 GB of RAM, and an NVIDIA GeForce 5060 Ti GPU 16 GB VRAM. The deep learning models were implemented using TensorFlow v2.20-dev.

4.2 Datasets

To ensure a comprehensive and robust evaluation, three IoMT datasets were chosen, each offering a unique challenge in terms of data structure and attack representation:

- **CICIoMT2024** [19]: developed by the Canadian Institute for Cybersecurity (CIC), it consists of 18 different cyberattacks. A key strength of this dataset is its protocol heterogeneity, making it highly representative of current IoMT environments.
- **IoMT-TrafficData (2023)** [11]: It encompasses benign traffic generated by IoMT protocols like CoAP, MQTT alongside malicious traffic from 8 distinct attack types. Selected for its specific IDS focus, This study utilized its granular packet-based view.
- **WUSTL-EHMS 2020** [9]: From Washington University in St. Louis offers a unique perspective by combining network-level data with patient biometric data. Generated from a real-time Enhanced Healthcare Monitoring System (EHMS).The dataset focuses on Man-in-the-Middle (MitM) attacks, specifically spoofing and data injection.

Table 1. Class Distribution for the Considered Datasets

Dataset: CICIoMT2024 (19 Classes)				Dataset: IoMT-TrafficData (2023) (9 Classes)				Dataset: WUSTL-EHMS 2020 (2 Classes)			
#	Class/Attack Type	Samples	%	#	Class/Attack Type	Samples	%	#	Class/Attack Type	Samples	%
0	Spoofing ARP	17,791	0.20%	0	Normal	148,812	43.7%	0	Normal	14,278	87.50%
1	Benign	230,339	2.63%	1	Apache-Killer	24,000	7.0%	1	Attack (Spoofing + Alteration)	2,040	12.50%
2	MQTT DDoS connect flood	214,952	2.45%	2	RUDY	24,000	7.0%				
3	MQTT DDoS publish flood	36,039	0.41%	3	Slow-Read	24,000	7.0%				
4	MQTT DoS publish flood	52,881	0.60%	4	Slow-Loris	24,000	7.0%				
5	MQTT DoS connect flood	15,904	0.18%	5	ARP-Spoofing	24,000	7.0%				
6	MQTT Malformed data	6,877	0.08%	6	CAM-Overflow	24,000	7.0%				
7	Recon OS scan	20,666	0.24%	7	MQTT-Malaria	24,000	7.0%				
8	Recon Ping sweep	926	0.01%	8	Net-Scan	24,000	7.0%				
9	Recon Port scan	106,603	1.22%								
10	Recon VulScan	3,207	0.04%								
11	DDoS ICMP	1,887,175	21.51%								
12	DDoS SYN	974,359	11.10%								
13	DDoS TCP	987,063	11.25%								
14	DDoS UDP	1,998,026	22.77%								
15	DoS ICMP	514,724	5.87%								
16	DoS SYN	540,498	6.16%								
17	DoS TCP	462,480	5.27%								
18	DoS UDP	704,503	8.03%								

4.3 Data Partitioning and Validation Strategy

Evaluation protocols were aligned with dataset-specific conventions. CICIoMT2024 and IoMT-TrafficData were tested on both original and binary distributions to ensure comparability with prior studies [7,31]. Similarly, WUSTL-EHMS 2020 was converted to binary by merging attack classes, following established approaches [18,32,43]. Regarding validation, IoMT-TrafficData and WUSTL-EHMS 2020 employed Stratified 5-Fold Cross-Validation. Conversely, CICIoMT2024 utilized the original authors' fixed partitions [3,33,43] to prioritize direct benchmarking against existing literature over k-fold validation.

4.4 Hyperparameter Configuration

To ensure the reproducibility of all experiments, a global random seed was consistently applied across all model initializations and data partitioning. Our first evaluation compares four model families: a standard Decision Tree (DT), a baseline Neural Network (NN), and two NNs leveraging our augmented datasets (*Binned Prediction* and *Binned Thresholding*). The number of top features to be selected has been chosen as 20. This value ($\approx 50\%$ of the feature space of the datasets) is a pragmatic balance between model representativeness and computational cost. Literature [32,53] shows that this top quantile of features typically captures the majority of a dataset's predictive signal. Crucially, sensitivity analysis confirms this 20-feature threshold as the optimal trade-off: incorporating additional features yields negligible predictive gains while linearly inflating the memory footprint, resulting in an unjustified waste of storage and computational resources. All neural network models share a common textbook configuration 2. The number of epochs was selected in accordance with the observations in

[32,37,43], which show that this value is sufficient to achieve maximum performance; furthermore, this guarantees an overall reduction in time, which is a key factor in this study. The batch size was tailored to the dataset: 256 for the larger CICIoMT2024 and IoMT-TrafficData datasets, and 64 for the smaller WUSTL-EHMS 2020 dataset. Each hidden layer utilizes ReLU activation followed by Batch Normalization to stabilize training. The output layer employs Softmax or Sigmoid for multi-class and binary tasks, respectively. The Decision Tree was tuned over its `max_leaf_nodes` and `max_depth`, while the Binned NN variants include an additional hyperparameter, α, which controls the granularity of the binning process. The specific values explored for this comprehensive hyperparameter space are summarized in Table 2.

Table 2. Hyperparameter Space for All Models (Global Seed: `42`)

Family	Param	Values/Settings
DT	Max Leaves	{5, 10, 20, 50, 100, 200, 500, None}
	Max Depth	{5, 10, 15, None}
NN	*Fixed Params*	**Arch**: [128, 64, 32]; **BN**: Yes after each layer; **Epochs**: 20; **LR**: 1e-3
	Batch Size	**256** (CICIoMT, TrafficData), **64** (WUSTL 2020)
	Binned Models Specifics:	
	Binning α	{50, 100, 200, 500}
	Top-feat k	20

4.5 Results and Discussion

This section presents the performance evaluation of the models across the three datasets. The datasets were tested using the classes shown in Table 1.

As illustrated in the comparison graphs (Figs. 2, 3, and 4), which show the performance of the best configuration for each dataset and model, it can be affirmed that the hybrid approach improves overall performance. Our models demonstrated top-tier results across all three datasets, outperforming the Decision Tree and Neural Network that compose them. Specifically, BinThresh prevails in the CICIoMT2024 dataset, while BinPred is the highest performing model for the WUSTL-EHMS 2020 dataset. In the IoMT-TrafficData dataset, both BinPred and BinThresh rank first in a tie; however, their performance is comparable to that of the Decision Tree, surpassing it by only 0.001 in Balanced Accuracy.

Table 3 provides a comparison between the results obtained from this study and those from the current literature. Missing metrics are not stated in the listed studies: for this reason, no comparisons are proposed for these metrics; however, these are reported for greater transparency. To facilitate comparison with other studies, Table 3 also shows the values for an additional experiment, i.e., the testing of our models in binary attack/benign classification on CICIoMT2024 and IoMT-TrafficData. In this case, the values shown are those that correspond to

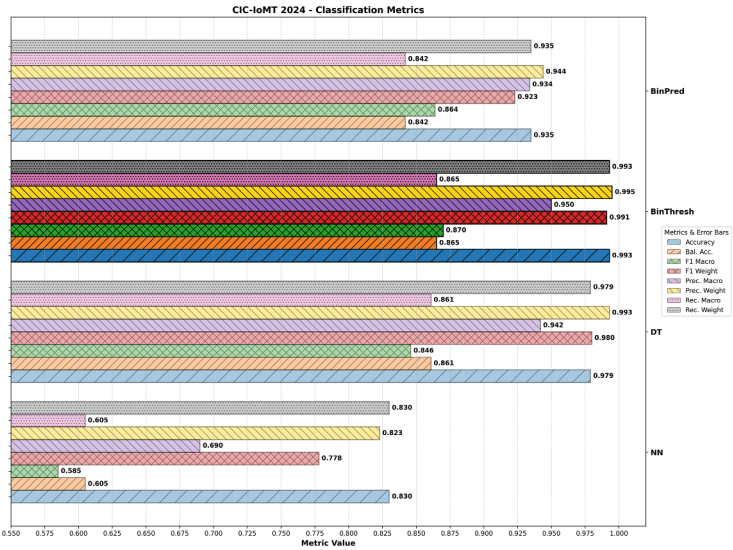


Fig. 2. Performance comparison between models (CICIoMT2024). For each metric, the value corresponding to the configuration that yields the highest result for that metric is reported. The variability in the bars refers to the variability observed between the folds in the same configuration.

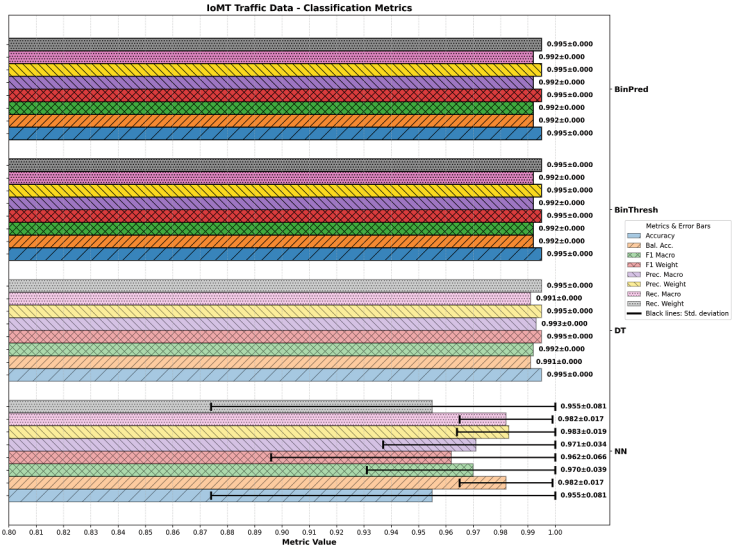


Fig. 3. Performance comparison between models (IoMT-TrafficData). For each metric, the value corresponding to the configuration that yields the highest result for that metric is reported. The variability in the bars refers to the variability observed between the folds in the same configuration.

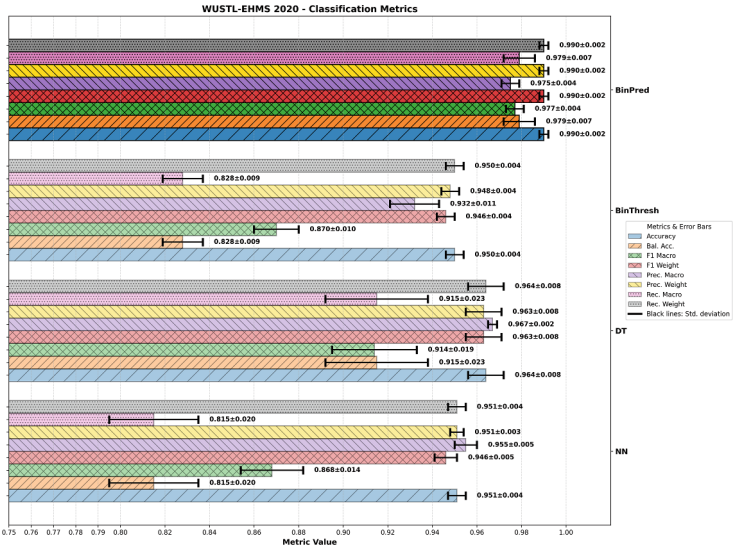


Fig. 4. Performance comparison between models (WUSTL-EHMS 2020). For each metric, the value corresponding to the configuration that yields the highest result for that metric is reported. The variability in the bars refers to the variability observed between the folds in the same configuration.

the model that achieves the best performance. For the CICIoMT2024 dataset, in the multi-class test, our approach achieves the best F1 Weight, Precision Weight, and Recall Weight metrics. Compared to the second-ranked CNN [43], it achieves an advantage of 0.011 in F1 Weight, 0.013 in Accuracy and Recall Weight, and 0.015 in Precision Weight. In accuracy, however, it is surpassed by [37] by 0.001. This is a perfectly acceptable margin considering that our approach is lighter in inference, thus allowing much more data to be processed in the IoMT scenario, where time is a critical factor. In the CICIoMT2024 test for binary classification, the best-performing model (BinThresh) achieves a remarkable 0.998 on F1 Weight, Precision Weight, Recall Weight, and Accuracy, surpassing the second-ranked model [7] by 0.003. It is meaningful that, both in the multi-class experiment and in the binary classification experiment, our approach outperforms computationally heavier models in inference [3,33,43]. Multi-class tests on the IoMT-TrafficData dataset show that the dominant model is the heavyweight XBiDeep [53], while our BinPred ranks second with a slight difference of 0.003 on Recall Weight and 0.004 on the other metrics. Given the minor difference in metrics, the lightweight nature of our model makes it an excellent solution in a real-world e-health context, offering an excellent trade-off between inference time and classification performance (0.995 across all compared metrics). It is also worth noting the great stability between folds. In the binary classification for IoMT-TrafficData, on the other hand, the proposed BinPred approach dominates all metrics, achieving an impressive 0.999 ± 0.000, outperforming other approaches

in terms of performance and lightweight nature. The tests for WUSTL-EHMS 2020 show more variable results. For F1 Weight, XAI-XGBoost [32] achieves the highest value with 0.991; however, our BinPred achieves a value of 0.990 ± 0.002, thus showing that there are folds where it outperforms XAI-XGBoost. Similar considerations apply to Precision Weight, where BinPred ranks first with a value of 0.990, tied with [16]; our variability leads it to exceed it by 0.002 in some folds. On Accuracy, our model achieves a score of 0.990 ± 0.002, which is slightly lower than the 0.992 achieved by XAI-XGBoost. XAI-XGBoost, on the other hand, dominates on Recall Weight with a score of 0.999. Even in this case, however, our model achieves excellent performance of at least 0.990 on all metrics in comparison, and is much lighter than [16,32], once again proving to be an excellent trade-off.

Table 3. Comprehensive Comparison of AI Models for Attack Detection in IoMT Environments

Study	Model	Inference Weight	Inf. Time (μs/samp)	F1 Weight	Prec. Weight	Rec. Weight	Acc.	F1 Macro	Bal. Acc.
Dataset: CICIoMT2024									
Multi-class Classification (19 Classes)									
L2D2 [3]	LSTM (2xLSTM + Dense)	Heavy	—	0.980	0.980	0.980	0.980	—	—
CNN-Based Detection [43]	CNN	Medium	—	0.980	0.980	0.990	0.990	—	—
InfoGain + MultiD-CNN [33]	MultiD-CNN (with IG)	Medium	—	0.979	0.977	0.982	0.981	—	—
Multi-view FL [37]	TabNet & MLP (Ensemble, FL)	Medium	—	—	—	—	**0.994**	—	—
CICIoMT2024 [19]	RF	Light-Medium	—	0.551	0.691	0.577	0.733	—	—
CICIoMT2024 [19]	LR	Light	—	0.432	0.547	0.471	0.727	—	—
CICIoMT2024 [19]	AdaBoost	Light	—	0.141	0.144	0.238	0.422	—	—
CICIoMT2024 [19]	DNN	Light	—	0.522	0.649	0.553	0.729	—	—
Our Study (BinThresh)	**Small NN + Bin Thresholding**	**Light**	9.0	**0.991**	**0.995**	**0.993**	0.993	0.870	0.865
Binary Classification (2 Classes)									
DDoSViT [2]	Vision Transformer (ViT)	Heavy	—	0.995	0.995	0.995	0.995	—	—
CNN-Based Detection [43]	CNN	Medium	—	0.990	0.990	0.990	0.990	—	—
Nilsimsa Fingerprinting [31]	RF (Nilsimsa)	Light-Medium	—	0.986	0.994	0.978	0.982	—	—
CICIoMT2024 [19]	RF	Light-Medium	—	0.961	0.971	0.951	0.996	—	—
CICIoMT2024 [19]	LR	Light	—	0.946	0.952	0.940	0.995	—	—
CICIoMT2024 [19]	AdaBoost	Light	—	0.959	0.959	0.961	0.996	—	—
CICIoMT2024 [19]	DeepNeuralNetwork	Light	—	0.952	0.956	0.948	0.996	—	—
Our Study (BinThresh)	**Small NN + Bin Thresholding**	**Light**	9.0	**0.998**	**0.998**	**0.998**	**0.998**	0.978	0.975
Dataset: IoMT-TrafficData									
Multi-class Classification (9 Classes)									
XBiDeep [53]	BI-GRU BI-LSTM	Heavy	—	**0.999**	**0.999**	**0.998**	**0.999**	—	—
IoMT-TrafficData [11]	RF	Light-Medium	—	0.994	0.994	0.994	0.994	—	—
IoMT-TrafficData [11]	DT	Light	—	0.951	0.970	0.938	0.938	—	—
IoMT-TrafficData [11]	LogR	Light	—	0.924	0.934	0.915	0.915	—	—
IoMT-TrafficData [11]	NB	Light	—	0.830	0.934	0.796	0.796	—	—
IoMT-TrafficData [11]	SVM	Light	—	0.918	0.922	0.915	0.915	—	—
IoMT-TrafficData [11]	DNN	Light	—	0.990	0.990	0.990	0.990	—	—
Our Study (BinPred)	**Small NN + Bin Prediction**	**Light**	7.9	$0.995_{\pm 0.000}$	$0.995_{\pm 0.000}$	$0.995_{\pm 0.000}$	$0.995_{\pm 0.000}$	$0.992_{\pm 0.000}$	$0.992_{\pm 0.000}$
Binary Classification (2 Classes)									
IoMT-TrafficData [11]	RF	Light-Medium	—	0.994	0.994	0.994	0.994	—	—
IoMT-TrafficData [11]	DT	Light	—	0.951	0.970	0.938	0.938	—	—
IoMT-TrafficData [11]	LogR	Light	—	0.924	0.934	0.915	0.915	—	—
IoMT-TrafficData [11]	NB	Light	—	0.830	0.934	0.796	0.796	—	—
IoMT-TrafficData [11]	SVM	Light	—	0.918	0.922	0.915	0.915	—	—
IoMT-TrafficData [11]	DNN	Light	—	0.990	0.990	0.990	0.990	—	—
Our Study (BinPred)	**Small NN + Bin Prediction**	**Light**	7.9	$\mathbf{0.999}_{\pm 0.000}$	$\mathbf{0.999}_{\pm 0.000}$	$\mathbf{0.999}_{\pm 0.000}$	$\mathbf{0.999}_{\pm 0.000}$	$0.999_{\pm 0.000}$	$0.999_{\pm 0.000}$
Dataset: WUSTL-EHMS 2020									
Binary Classification (2 Classes)									
PSO-DNN [18]	PSO-DNN	Heavy	—	0.950	0.960	0.960	0.960	—	—
Mutual Information FS [5]	LRGU-MIFS	Heavy	—	—	—	—	0.949	—	—
Towards Robust IoMT Models [16]	Deep Sequential Model	Heavy	—	0.990	**0.990**	0.990	0.990	—	—
XAI-XGBoost [32]	XGBoost (SMOTE-ENN)	Medium-Heavy	—	**0.991**	0.984	**0.999**	**0.992**	**0.990**	**0.990**
RFE-MLP [38]	RFE-MLP	Medium	—	0.962	0.962	0.962	0.962	—	—
A Tree Classifier Based IDS [28]	Robust Scaler + RF + Data augmentation	Light	3.5	0.938	0.935	0.909	0.942	—	—
RFE/Ridge [40]	RFE-DT	Light	—	0.911	0.965	0.863	0.979	—	—
IDS Comparison [29]	SVM	Light	—	—	—	—	0.924	—	—
Security Analysis [34]	Ensemble (LR + KNN)	Light	—	0.607	0.965	0.444	0.925	—	—
Our Study (BinPred)	**Small NN + Bin Prediction**	**Light**	28.6	$\mathbf{0.990}_{\pm 0.002}$	$\mathbf{0.990}_{\pm 0.002}$	$0.990_{\pm 0.002}$	$\mathbf{0.990}_{\pm 0.002}$	$0.977_{\pm 0.004}$	$0.979_{\pm 0.007}$

The inference time on the CICIoMT2024 and IoMT-TrafficData datasets is 9.0 and 7.9 μs/samp, respectively; this implies that they are capable of processing 111,000 and 126,582 examples per second, respectively. On the WUSTL-EHMS dataset, the times were longer (28.6 μs/samp). The remarkably low inference times measured for our model substantiate its suitability for real-time deployment in resource-constrained IoMT environments. It is important to contextualize this performance: while enterprise-grade IDS/IPS may prioritize terabit-scale throughput on network backbones, the defining challenge in

IoMT is achieving high-accuracy detection under strict computational limits. Our solution meets this challenge directly. With a throughput of over 34,000 classifications per second on the WUSTL-EHMS dataset and over 110,000 for the other two datasets, the system possesses orders of magnitude more capacity than required to process the predictable, low-volume flow data generated by a typical hospital ward of IoMT devices. Beyond inference efficiency, our model's minimal footprint ensures sustainable lifecycle costs and enables local retraining on hospital servers. While supervised binning mitigates overfitting by preventing noise memorization, detecting true zero-day anomalies remains challenging. Therefore, assessing robustness against unprecedented attacks requires future validation on continuous, real-world data streams. In conclusion, the proposed approach represents an optimal compromise between speed and classification performance.

5 Conclusions

This paper proposed a lightweight hybrid IDS/IPS approach that combines a decision tree model with a compact neural network architecture. The proposed methodology is driven by the insight that tree-based models excel at handling non-continuous and segmented feature spaces common in IoMT data, while neural networks can complement them by capturing subtle non-linear relationships. Through the introduction of a supervised binning process for feature discretisation, the approach improves robustness and computational efficiency without sacrificing detection accuracy. Experimental validation on three reference datasets (CICIoMT2024, IoMT-TrafficData, and WUSTL-EHMS 2020) confirmed the effectiveness of the proposed approach. In most cases, the hybrid methods (BinPred and BinThresh) outperformed or matched state-of-the-art, resource-intensive deep learning solutions while delivering inference times that fully support real-time operation. Specifically, our models achieved weighted F1 scores and accuracy scores above 0.99 in all the datasets, with the ability to process from 30,000 to 100,000 examples per second. These results demonstrate that lightweight models can achieve high-level performance while remaining usable for deployment on embedded or edge medical devices. In experiments where our methodology did not rank first in the metrics, it still achieved second place behind significantly heavier models. Tests conducted in both multiclass and binary classification showed the strength of the model in both recognising and classifying attacks in IoMT networks. While theoretically suitable for edge deployment, a current limitation is the lack of validation on physical IoT/IoMT hardware (e.g., ARM microcontrollers). Standard environments were used strictly to ensure benchmark comparability. Future work will prioritize on-device testing to evaluate real-world power consumption, thermal behavior, and exact hardware latency within a live hospital testbed. Furthermore, the variance calculated on the metrics showed the particular robustness of the model. In conclusion, our work shows how a solution that exploits the peculiarities of IoMT networks can be employed for the relative requirements of speed and accuracy in attack detection.

Acknowledgments. – **Funding:** This research was co-funded by the Italian Complementary National Plan PNC-I.1 "Research initiatives for innovative technologies and pathways in the health and welfare sector" D.D. 931 of 06/06/2022, "DARE - DigitAl lifelong pRevEntion" initiative, code PNC0000002, CUP: B53C22006450001.

– **Use of Generative AI:** The authors declare that no generative AI was used in the writing of the paper.

– **Data availability:** Data will be made available on request.

Disclosure of Interests. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Annual number of IoT attacks global 2022 | Statista — statista.com. https://www.statista.com/statistics/1377569/worldwide-annualinternet-of-things-attacks/. Accessed 11 June 06 2025
2. DecisionTreeDiscretiser — 1.9.3 — feature-engine.trainindata.com. https://feature-engine.trainindata.com/en/latest/user_guide/discretisation/DecisionTreeDiscretiser.html. Accessed 11 July 2025
3. Akar, G., Sahmoud, S., Onat, M., Cavusoglu, Ü., Malondo, E.: L2D2: a novel LSTM model for multi-class intrusion detection systems in the era of IOMT. IEEE Access (2025)
4. Al-Turjman, F., Nawaz, M.H., Ulusar, U.D.: Intelligence in the internet of medical things era: a systematic review of current and future trends. Comput. Commun. **150**, 644–660 (2020)
5. Alalhareth, M., Hong, S.C.: An improved mutual information feature selection technique for intrusion detection systems in the internet of medical things. Sensors **23**(10), 4971 (2023)
6. Alharthi, A., Alaryani, M., Kaddoura, S.: A comparative study of machine learning and deep learning models in binary and multiclass classification for intrusion detection systems. Array, 100406 (2025)
7. Ali, M., Saleem, Y., Hina, S., Shah, G.A.: DDOSVIT: IoT DDOS attack detection for fortifying firmware over-the-air (OTA) updates using vision transformer. Internet Things **30**, 101527 (2025)
8. Alsaedi, A., Moustafa, N., Tari, Z., Mahmood, A., Anwar, A.: Ton_IoT telemetry dataset: a new generation dataset of IoTand IIoT for data-driven intrusion detection systems. IEEE Access **8**, 165130–165150 (2020)
9. Alsalman, D.: A comparative study of anomaly detection techniques for IoT security using adaptive machine learning for IoT threats. IEEE Access **12**, 14719–14730 (2024)
10. Alsubaei, F., Abuhussein, A., Shandilya, V., Shiva, S.: IOMT-SAF: internet of medical things security assessment framework. Internet Things **8**, 100123 (2019)
11. Areia, J., Bispo, I.A., Santos, L., Costa, R.L.D.C.: IOMT-trafficdata: dataset and tools for benchmarking intrusion detection in internet of medical things. IEEE Access **12**, 115370–115385 (2024)
12. Aversano, L., Bernardi, M.L., Cimitile, M., Montano, D., Pecori, R., Veltri, L.: Explainable anomaly detection of synthetic medical IoT traffic using machine learning. SN Comput. Sci. **5**(5), 488 (2024)

13. Baho, S.A., Abawajy, J.: Analysis of consumer IoT device vulnerability quantification frameworks. Electronics **12**(5), 1176 (2023)
14. Benkhelifa, E., Welsh, T., Hamouda, W.: A critical review of practices and challenges in intrusion detection systems for IoT: toward universal and resilient systems. IEEE Commun. Surv. Tutorials **20**(4), 3496–3509 (2018)
15. Bhushan, B., Kumar, A., Agarwal, A.K., Kumar, A., Bhattacharya, P., Kumar, A.: Towards a secure and sustainable internet of medical things (IOMT): requirements, design challenges, security techniques, and future trends. Sustainability **15**(7), 6177 (2023)
16. Bouke, M.A., El Atigh, H., Abdullah, A.: Towards robust and efficient intrusion detection in IOMT: a deep learning approach addressing data leakage and enhancing model generalizability. In: Multimedia Tools and Applications, pp. 1–20 (2024)
17. Chaabouni, N., Mosbah, M., Zemmari, A., Sauvignac, C., Faruki, P.: Network intrusion detection for IoT security based on learning techniques. IEEE Commun. Surv. Tutorials **21**(3), 2671–2701 (2019)
18. Chaganti, R., Mourade, A., Ravi, V., Vemprala, N., Dua, A., Bhushan, B.: A particle swarm optimization and deep learning approach for intrusion detection system in internet of medical things. Sustainability **14**(19), 12828 (2022)
19. Dadkhah, S., Neto, E.C.P., Ferreira, R., Molokwu, R.C., Sadeghi, S., Ghorbani, A.A.: CICIOMT 2024: a benchmark dataset for multi-protocol security assessment in IOMT. Internet Things **28**, 101351 (2024)
20. Dentamaro, V., Convertini, V.N., Galantucci, S., Giglio, P., Palmisano, T., Pirlo, G.: Ensemble consensus: an unsupervised algorithm for anomaly detection in network security data. ITASEC, pp. 309–318 (2021)
21. Elnakib, O., Shaaban, E., Mahmoud, M., Emara, K.: EIDM: deep learning model for IoT intrusion detection systems. J. Supercomput. **79**(12) (2023)
22. Eskandari, M., Janjua, Z.H., Vecchio, M., Antonelli, F.: PassBan ids: an intelligent anomaly-based intrusion detection system for IoT edge devices. IEEE Internet Things J. **7**(8), 6882–6897 (2020)
23. Faruqui, N., et al.: Safetymed: a novel IOMT intrusion detection system using CNN-LSTM hybridization. Electronics **12**(17), 3541 (2023)
24. Garcia, S., Luengo, J., Sáez, J.A., Lopez, V., Herrera, F.: A survey of discretization techniques: taxonomy and empirical analysis in supervised learning. IEEE Trans. Knowl. Data Eng. **25**(4), 734–750 (2012)
25. Ghazanfar, S., Hussain, F., Rehman, A.U., Fayyaz, U.U., Shahzad, F., Shah, G.A.: IoT-flock: an open-source framework for IoT traffic generation. In: 2020 International Conference on Emerging Trends in Smart Technologies (ICETST), pp. 1–6. IEEE (2020)
26. Ghubaish, A., Salman, T., Zolanvari, M., Unal, D., Al-Ali, A., Jain, R.: Recent advances in the internet-of-medical-things (IOMT) systems security. IEEE Internet Things J. **8**(11), 8707–8718 (2020)
27. Guo, G., Pan, X., Liu, H., Li, F., Pei, L., Hu, K.: An IoT intrusion detection system based on ton IoT network dataset. In: 2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC), pp. 0333–0338. IEEE (2023)
28. Gupta, K., Sharma, D.K., Gupta, K.D., Kumar, A.: A tree classifier based network intrusion detection model for internet of medical things. Comput. Electr. Eng. **102**, 108158 (2022)
29. Hady, A.A., Ghubaish, A., Salman, T., Unal, D., Jain, R.: Intrusion detection system for healthcare systems using medical and network data: a comparison study. IEEE Access **8**, 106576–106584 (2020)

30. Hathaliya, J.J., Tanwar, S.: An exhaustive survey on security and privacy issues in healthcare 4.0. Comput. Commun. **153**, 311–335 (2020)
31. Hernandez-Jaimes, M.L., Martínez-Cruz, A., Ramírez-Gutiérrez, K.A., Guevara-Martínez, E.: Enhancing machine learning approach based on nilsimsa fingerprinting for ransomware detection in IOMT. IEEE Access (2024)
32. Hosain, Y., Çakmak, M.: XAI-XGBOOST: an innovative explainable intrusion detection approach for securing internet of medical things systems. Sci. Rep. **15**(1), 22278 (2025)
33. Hossain, M.T., Meem, S.M., Setu, J.H., Halder, N., Islam, A., Alam, M.Z.: Cyber-attacks classification on internet of medical things using information gain feature selection and machine learning. In: 2024 Advances in Science and Engineering Technology International Conferences (ASET), pp. 1–10. IEEE (2024)
34. Ibrahim, M., Al-Wadi, A., Elhafiz, R.: Security analysis for smart healthcare systems. Sensors **24**(11), 3375 (2024)
35. Jacod, J.: Discretization of Processes, Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-24127-7
36. Jayalaxmi, P., Saha, R., Kumar, G., Alazab, M., Conti, M., Cheng, X.: Pignus: a deep learning model for ids in industrial internet-of-things. Comput. Secur. **132**, 103315 (2023)
37. Jeremiah, S.R., El Azzaoui, A., Gritzalis, S., Park, J.H.: Multi-view learning and model fusion framework for threat detection in multi-protocol IOMT networks. Inform. Fusion, 103435 (2025)
38. Kilincer, I.F., Ertam, F., Sengur, A., Tan, R.S., Acharya, U.R.: Automated detection of cybersecurity attacks in healthcare systems with recursive feature elimination and multilayer perceptron optimization. Biocybernet. Biomed. Eng. **43**(1), 30–41 (2023)
39. Kotronis, C., et al.: Evaluating internet of medical things (IOMT)-based systems from a human-centric perspective. Internet Things **8**, 100125 (2019)
40. Lazrek, G., Chetioui, K., Balboul, Y., Mazer, S., et al.: An RFE/RIDGE-ML/DL based anomaly intrusion detection approach for securing IOMT system. Results Eng. **23**, 102659 (2024)
41. Liu, H., Hussain, F., Tan, C.L., Dash, M.: Discretization: an enabling technique. Data Min. Knowl. Disc. **6**(4), 393–423 (2002)
42. Mirlashari, M., Rizvi, S.A.M.: Enhancing IoT intrusion detection system with modified e-graphsage: a graph neural network approach. Int. J. Inf. Technol. **16**(4), 2705–2713 (2024)
43. Mohammadi, A., Ghahramani, H., Asghari, S.A., Aminian, M.: Securing healthcare with deep learning: a CNN-based model for medical IoT threat detection. In: 2024 19th Iranian Conference on Intelligent Systems (ICIS), pp. 168–173. IEEE (2024)
44. Mukhtar, B.I., Elsayed, M.S., Jurcut, A.D., Azer, M.A.: IoT vulnerabilities and attacks: silex malware case study. Symmetry **15**(11), 1978 (2023)
45. Musthafa, M.B., et al.: Optimizing IoT intrusion detection using balanced class distribution, feature selection, and ensemble machine learning techniques. Sensors **24**(13), 4293 (2024)
46. Naeem, H., Alsirhani, A., Alserhani, F., Ullah, F., Krejcar, O.: Augmenting internet of medical things security: deep ensemble integration and methodological fusion. Comput. Model. Eng. Sci. **141**(3), 2185 (2024)
47. Naghib, A., Gharehchopogh, F.S., Zamanifar, A.: A comprehensive and systematic literature review on intrusion detection systems in the internet of medical things: current status, challenges, and opportunities. Artif. Intell. Rev. **58**(4), 114 (2025)

48. Neto, E.C.P., Dadkhah, S., Sadeghi, S., Molyneaux, H., Ghorbani, A.A.: A review of machine learning (ml)-based IoT security in healthcare: a dataset perspective. Comput. Commun. **213**, 61–77 (2024)
49. Nie, F., Liu, W., Liu, G., Gao, B.: M2VT-IDS: a multi-task multi-view learning architecture for designing IoT intrusion detection system. Internet Things **25**, 101102 (2024)
50. Seger, C.: An investigation of categorical variable encoding techniques in machine learning: binary versus one-hot and feature hashing (2018)
51. Shammar, E., Cui, X., Zahary, A., Alsamhi, S.H., Al-qaness, M.A.: Threat to trust: a systematic review on internet of medical things security. J. Parallel Distrib. Comput. **206**, 105172 (2025). https://doi.org/10.1016/j.jpdc.2025.105172, https://www.sciencedirect.com/science/article/pii/S074373152500139X
52. Sharafaldin, I., Habibi Lashkari, A., Ghorbani, A.A.: A detailed analysis of the cicids2017 data set. In: Mori, P., Furnell, S., Camp, O. (eds.) International Conference on Information Systems Security and Privacy. ICISSP 2018. Communications in Computer and Information Science, vol. 977, pp. 172–188. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-25109-3_9
53. Turgut, Z., Başarslan, M.S.: XBIDEEP: a novel explainable artificial intelligence based intrusion detection system for internet of medical things environment. Internet Things, 101675 (2025)
54. Vishnu, S., Ramson, S.J., Jegan, R.: Internet of medical things (IOMT)-an overview. In: 2020 5th International Conference on Devices, Circuits and Systems (ICDCS), pp. 101–104. IEEE (2020)
55. Wang, Z., Chen, H., Yang, S., Luo, X., Li, D., Wang, J.: A lightweight intrusion detection method for IoT based on deep learning and dynamic quantization. PeerJ Comput. Sci. **9**, e1569 (2023)
56. Wang, Z., Li, J., Yang, S., Luo, X., Li, D., Mahmoodi, S.: A lightweight IoT intrusion detection model based on improved Bert-of-Theseus. Expert Syst. Appl. **238**, 122045 (2024)
57. Yoo, J., Min, B., Kim, S., Shin, D., Shin, D.: Study on network intrusion detection method using discrete pre-processing method and convolution neural network. IEEE Access **9**, 142348–142361 (2021). https://doi.org/10.1109/ACCESS.2021.3120839
58. Zarpelão, B.B., Miani, R.S., Kawakani, C.T., De Alvarenga, S.C.: A survey of intrusion detection in internet of things. J. Netw. Comput. Appl. **84**, 25–37 (2017)
59. Zhong, M., Lin, M., Zhang, C., Xu, Z.: A survey on graph neural networks for intrusion detection systems: methods, trends and challenges. Comput. Secur. **141**, 103821 (2024)

Preserving Our Past: Spatial and Remote Sensing Technologies for Cultural Heritage in a Changing Climate (POP 2026)

From Open Land to Built Surfaces: Spatial Analysis and Remote Sensing for Photovoltaic Energy Computing

Maria Danese(✉) and Canio Alfieri Sabia

CNR-ISPC, Tito Scalo (PZ), Italy
maria.danese@cnr.it

Abstract. This work comes from a simple but urgent question: how can we produce clean energy without consuming our most precious resources, namely soil, landscapes and the memory of places?

Photovoltaics is one of the great hopes of the energy transition. But when it is installed on the ground, transforming countryside, hills and natural areas into fields of panels, we risk replacing one problem with another. The landscape changes, the soil is lost, and local communities perceive it as an imposition rather than a shared transformation.

Our contribution aims to change this point of view: what if all this energy could be generated without consuming new space, but by enhancing what already exists? The roofs of houses, agricultural sheds and industrial buildings can become a new "energy surface", capable of producing without taking away, thus minimizing the impact on the landscape and land consumption.

The research proposes an integrated method of spatial analysis and remote sensing that helps to design scenarios in which the ecological transition respects the meaning of places and does not pit energy against landscape, but allows them to coexist; in fact photovoltaics can only be sustainable if we start from a new perspective, capable of accompanying the use of technology.

Keywords: photovoltaic · landscape · spatial analysis · solar radiation · remote sensing

1 Introduction

The energy transition is a global challenge that is increasingly affecting public policy, with significant effects on territorial, economic and social systems. In Europe, with the "Clean Energy for All Europeans" package, the Commission has been promoting since 2016 an acceleration of the transition to clean energy by 2030, promoting an increase in the share of energy from renewable sources to 42.5% by that date [1]. Subsequently, Directive (EU) 2018/2001 on the promotion of the use of energy from renewable sources [2] and the Communication on the European Green Deal and Regulation (EU) 2021/1119 [3], establish the framework for achieving climate neutrality.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 371–384, 2027.
https://doi.org/10.1007/978-3-032-30530-5_23

The EU's commitment is therefore focused on reducing emissions and promoting sustainable growth models, which has led, especially in recent years, to a significant increase in the number of renewable energy plants.

However, while the spread of renewable energy sources is an essential step towards low-emission energy systems, the methods used to install the related infrastructure are not always without problems. The installation of large photovoltaic (PV) plants in rural areas or areas of environmental and landscape interest contributes to land consumption, alters the hydrogeological balance of the land and has a negative impact on the quality of the landscape and the protection of natural environments, raising questions about the real sustainability of such choices. The transformation of large rural areas into "photovoltaic fields" is perceived by local communities as an irreversible alteration and aesthetic degradation of the rural landscape, which is often protected for its historical and cultural value [4, 5].

With regard to the aforementioned land consumption, according to the European Green to Grey project (https://greentogrey.eu/), Europe loses approximately 1,500 km^2 of natural and agricultural areas every year. This data shows us how land transformations must be carefully evaluated and planned within the context of territorial and energy policies.

This project therefore aims to identify a methodological approach based on spatial analysis and remote sensing capable of estimating, once a large sample area has been identified, how much energy can be produced by converting the installation of panels on buildings and whether this energy is sufficient to cover the energy needs of the study areas. Consequently, energy modelling and scenario assessments are carried out, allowing energy suitability maps to be set up and the real contribution of PV on buildings to the decarbonisation of the local energy system to be assesses.

The expected results will provide evidence-based recommendations to support policy makers, spatial planners and stakeholders, promoting a more balanced integration of renewable energy, land use planning, agricultural landscape conservation and cultural heritage protection.

2 Study Area

The study area includes the three neighboring municipalities (Fig. 1) of Tito, Sant'Angelo Le Fratte and Satriano di Lucania, located in the Basilicata Region (southern Italy) and treated as a single case study, as they have similar geomorphological, settlement and climatic conditions. In particular, with regard to settlement, all three municipalities have historic centers with a strong landscape identity and the simultaneous presence of an industrial and craft area, which is particularly developed in Tito. In addition, all three municipalities have fairly active agricultural areas.

The sample chosen is representative of many rural areas in inland Italy, meaning that the proposed framework can easily be transferred to other similar contexts.

A preliminary mapping of the PV panels present in the studied area shows the results reported in the Tables 1 and 2 and represented in Figs. 2 and 3: in Tito, PV is distributed in a complex mix, energy-intensive but land-consuming, as a large number of plants are located on industrial land, but also on agricultural land. In Satriano di Lucania, there is

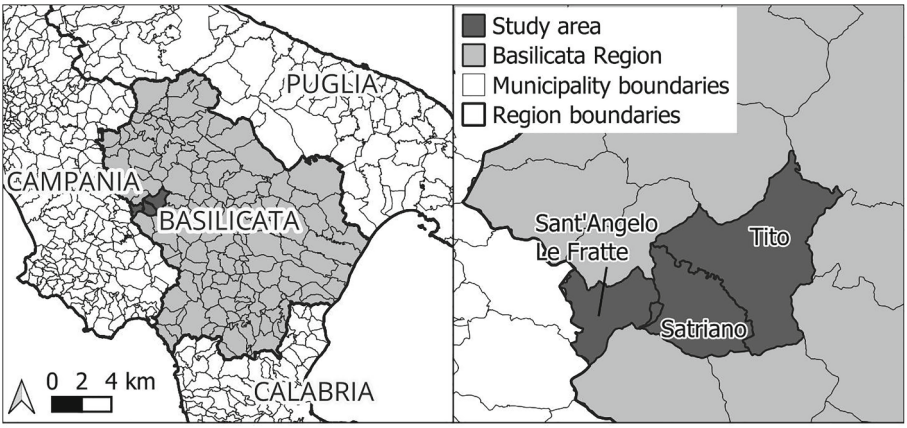


Fig. 1. Territorial overview of the study area.

strong pressure on the land as PV is mainly ground-mounted. In Sant'Angelo Le Fratte, mainly already built-up areas have been exploited, so at present, land consumption by PV plants is practically zero.

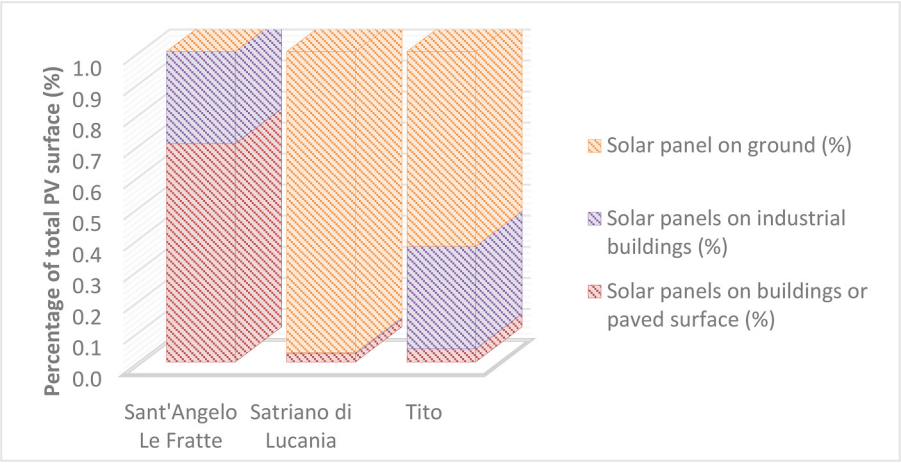


Fig. 2. Distribution (%) of PV installations by support type in the three studied Municipalities.

In addition, a building density was calculated for the study area (Fig. 4). The upper box shows the density taking into account only buildings, while the lower box also includes PV panels located in the area. The areas outlined in black show how the construction of these plants leads to an increase in density and, consequently, to land consumption.

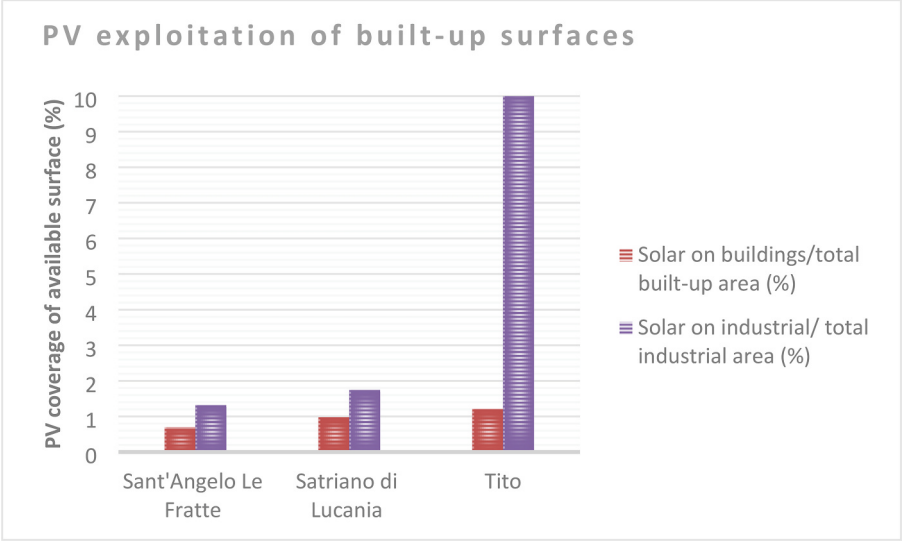


Fig. 3. PV exploitation (%) of built-up surfaces (not-industrial in red vs industrial in violet).

Table 1. PV areas extension (m^2) and percentage on buildings/paved, industrial elements and ground-mounted for the municipalities studied.

Municipality	PV areas on buildings or paved surface (m^2)	PV areas on industrial buildings (m^2)	PV areas on ground (m^2)	Total PV areas (m^2)	% PV on buildings or paved surface (%)	% PV on industrial buildings	% PV on ground
Sant'Angelo Le Fratte	1118	472	0	1591	70.3	29.7	0.0
Satriano di Lucania	2696	433	99359	102488	2.6	0.4	96.9
Tito	7149	55966	106152	169267	4.2	33.1	62.7

3 Materials and Methods

3.1 Datasets Used in This Work

In this paper multi-source spatial datasets were used to spatial assess the comparison between PV suitability on built and agricultural surfaces. More in particular the following resources were used:

Table 2. Data related to PV areas distribution in not-industrial or industrial built-up areas.

Municipality	Built-up areas excluding industrial (m^2)	Industrial areas (m^2)	% PV areas in built-up areas	% PV areas in industrial areas
Sant'Angelo Le Fratte	161931	35436	0.69	1.33
Satriano di Lucania	275038	24861	0.98	1.74
Tito	585005	543865	1.22	10.29

- Digital Surface Models (DSM), provided by Basilicata Region, to consider elevation and including buildings and vegetation. From the DSM slope and aspect were derived, in order to calculate solar radiation in a more realistic way.
- Building footprints, from the Regional Spatial Data Infrastructure (RSDI) of Basilicata, also containing the use information for each polygon.
- Data about the features protected by the Regional Landscape Plan, from the Regional Spatial Data Infrastructure (RSDI) of Basilicata, also containing the needed information for each polygon.
- Cloud cover, provided by Copernicus Sentinel-2 MSI Level-2A Surface Reflectance, accessed via Google Earth Engine (GEE; collection: COPERNICUS/S2_SR). In particular cloud information was derived from the Scene Classification Layer (SCL).
- Population data were obtained from Istituto Nazionale di Statistica (ISTAT).

3.2 Google Earth Engine©

GEE is a free, powerful platform that allows the analysis and monitoring of remote sensing big data at the planetary scale.

The first GEE version went out in 2001, but it was publicly released in 2010. 2001 could be the internal release date of the platform. And today it is still used very much in many application fields [6] in general for the global-scale earth observation data extraction and analysis [7], but more in particular, just to cite some examples, for land use mapping [8] and monitoring and natural risk assessment ([9] and reference therein), for population mapping [10] and climate data analysis [11].

The use of GEE offers many advantages. It gives the possibility to work with large dimension data (satellite data but also vector GIS data), without downloading them on pc, even if the download is possible, also because many of these data are freely available. At the same time, the calculation time is very fast, thanks to the cloud elaboration, more than it would be possible on a personal computer. It also gives the possibility to upload the user's data, if needed. Moreover, the GEE platform has an efficient versioning system and offers a flexible access to its code through APIs. Finally, even if the code developed by the user is online, the ownership remains to its developer, even if it could decide to share it.

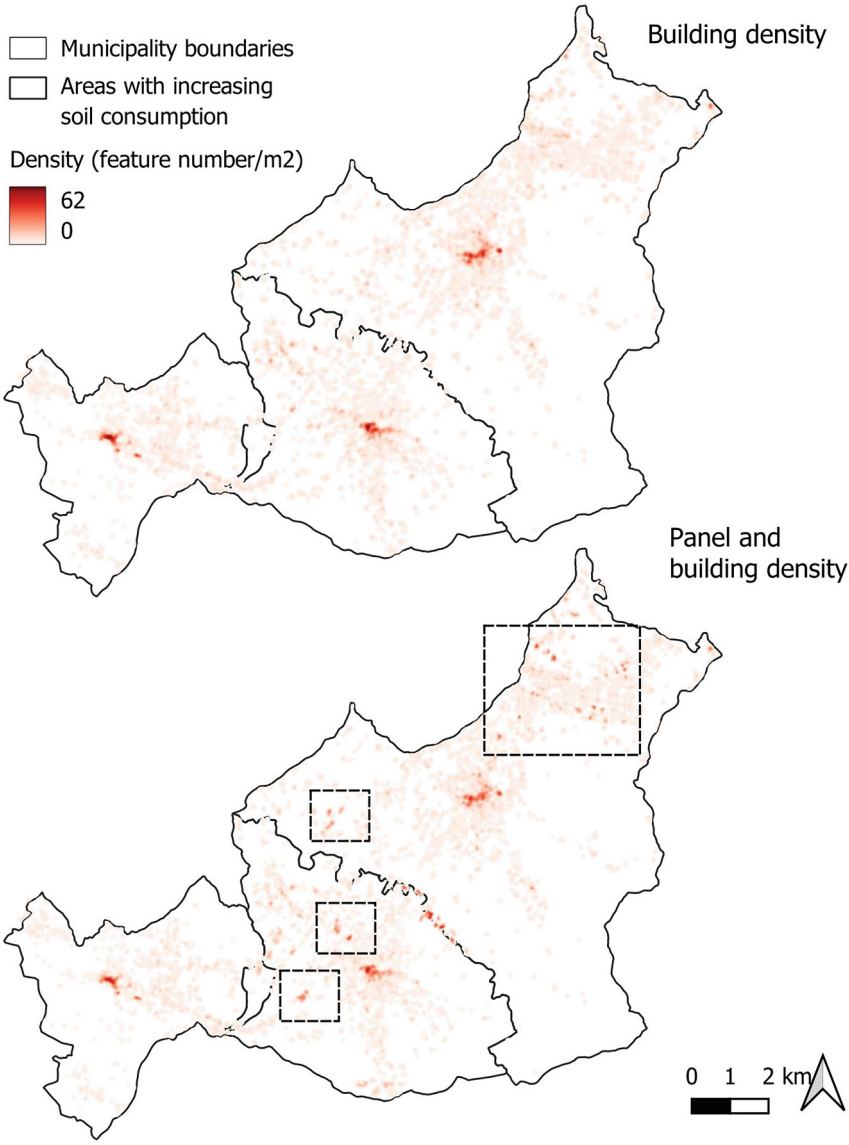


Fig. 4. Building and panel plus building density distribution.

3.3 Solar Radiation

The global solar radiation is the sum of the direct, diffuse and reflected radiation coming from the sun.

In this paper, for the calculation of the solar radiation direct and diffuse components, the Raster Solar Radiation tool [12–14] in ArcGIS Pro (ESRI) was used. It allows to map

the effects of the sun in a chosen period; in particular it allows to compute the incoming solar radiation, also called insolation, in every pixel of a raster surface.

To do this, the tool implements a multistep process:

1. It implements an upward-looking hemispherical viewshed, overlayed with a sky map expressed as azimuth angle useful to see how much of the sky is visible in the chosen time interval, by azimuth direction and for each angle of elevation; in this way a sort of horizon profile is then produced.
2. After this, the Direct Normal Irradiance can be calculated, because for the considered hours, days and latitude, it is possible to understand if from each pixel the sky (and consequently the sun) is visible or covered.
3. Then the Diffuse sky map is calculated. In this case the radiation derives from all the sky vault even if the local horizon is considered.
4. The first three steps are repeated for all the pixels, where the direct and diffuse radiation are summed for all the period considered.

The result is an insolation map expressed in kWh/m^2.

The main parameter that influences this calculation is topography; in fact, the spatial variation of elevation, slope, aspect and hillshade affects the total of insolation registered in each pixel. For this reason it is important to have: (1) high-resolution layers representing topography (2) to use, together with a Digital Elevation Model (DEM) a Digital Surface Model (DSM). This last, in fact, by reporting the presence of object on the terrain surfaces could change dramatically the hillshade distribution during the day and consequently also the cumulated insolation.

It also depends from time of the day (night vs day) and of the year (seasonality) and from the presence of clouds or dust, that scatters it.

4 Results

4.1 First Solar Radiation Calculation

Firstly the solar radiation was calculated for each season (Fig. 5). As expected, the seasonal solar radiation map varies depending on spatial variations in the morphology of the territory, increasing on south-facing slopes and ridge areas and decreasing on north-facing slopes and valley floor areas. The radiation classes represented were adapted to the range of each season, with an increase in cumulative radiation in spring and especially in summer, while it decreases in autumn and especially in winter.

4.2 Cloud Coverage Correction

At this point, in order to make a realistic estimate of the solar energy that can be produced in the study area, cloud cover must also be assessed. In fact, solar radiation is calculated in ArcGIS under unrealistic conditions of constant clear skies. For this purpose, a script has been developed in Google Earth Engine (GEE) that can estimate, with the help of Sentinel-2 Surface Reflectance data, the fraction of cloud cover on a daily basis using the Scene Classification Layer (SCL). This makes it possible to extrapolate, on

a daily basis, the average area fraction of the study area covered by medium and high probability cloud cover and cirrus clouds (8, 9 and 10). Of these bands: (1) 8 is more representative of widespread cloud cover involving moderate attenuation, (2) 9 reports the presence of very opaque clouds and therefore strong attenuation of direct radiation, but underestimates the presence of days with more extensive but less thick cover, 3) 10 takes into account cirrus clouds, which have a variable but generally less intense effect on radiation. Therefore, for prudential purposes and to obtain a summary indicator of cloud cover, the maximum value among the three daily fractions was calculated, thus considering the most penalising conditions.

The cloud cover values were then aggregated into four percentage classes (0–25%, 25–50%, 50–75% and 75–100%) on a seasonal basis for 2025, as shown in Table 3.

In order to calculate a seasonal weighted average based on the frequency of atmospheric conditions in relation to the type of sky, Cloud Reduction Factors (CRF) were introduced (Table 4). These are useful for attenuating solar radiation with clear skies (R_{cs}) and calculating a more realistic radiation value (R_{real}), according to Eq. 1.

$$R_{real} = R_{clear-sky} \cdot F_{season} \tag{1}$$

The correction factor (F_{season}) was calculated by combining the day number (d_i) and the CRF_i with a i cloud coverage percentage according to (Eq. 2):

$$F = \frac{d_{0-25}CRF_{0-25} + d_{25-50}CRF_{25-50} + d_{50-75}CRF_{50-75} + d_{75-100}CRF_{75-100}}{N} \tag{2}$$

where N is the total days of the season.

4.3 PV Surfaces Selection

After calculating solar radiation, buildings suitable for PV installation were selected, on the basis of which solar energy production was subsequently estimated.

The following criteria were used:

1. Buildings not belonging to the cultural heritage, including those located in the historic centers of the three municipalities in question.
2. Buildings with available surfaces of no less than $10m^2$.
3. Buildings with roofs with a slope of no more than 60°. This was calculated by estimating the median slope derived from the DSM.
4. Buildings with an industrial use all have a flat slope, so they were all considered.
5. Buildings with a residential use mainly have roofs with a gable roof. The direction of the ridge line was calculated. If the roof has a north and south orientation the half of the surfaces was considered for PV (only for the part exposed to south). Otherwise, with an east-south orientation the total surface was considered suitable for PV installation.

Furthermore, only free elements on which PVs are not already positioned were considered in the calculation, as already explained in Sect. 2. Totally 1,054,800 m^2 of suitable rooftop surfaces were found.

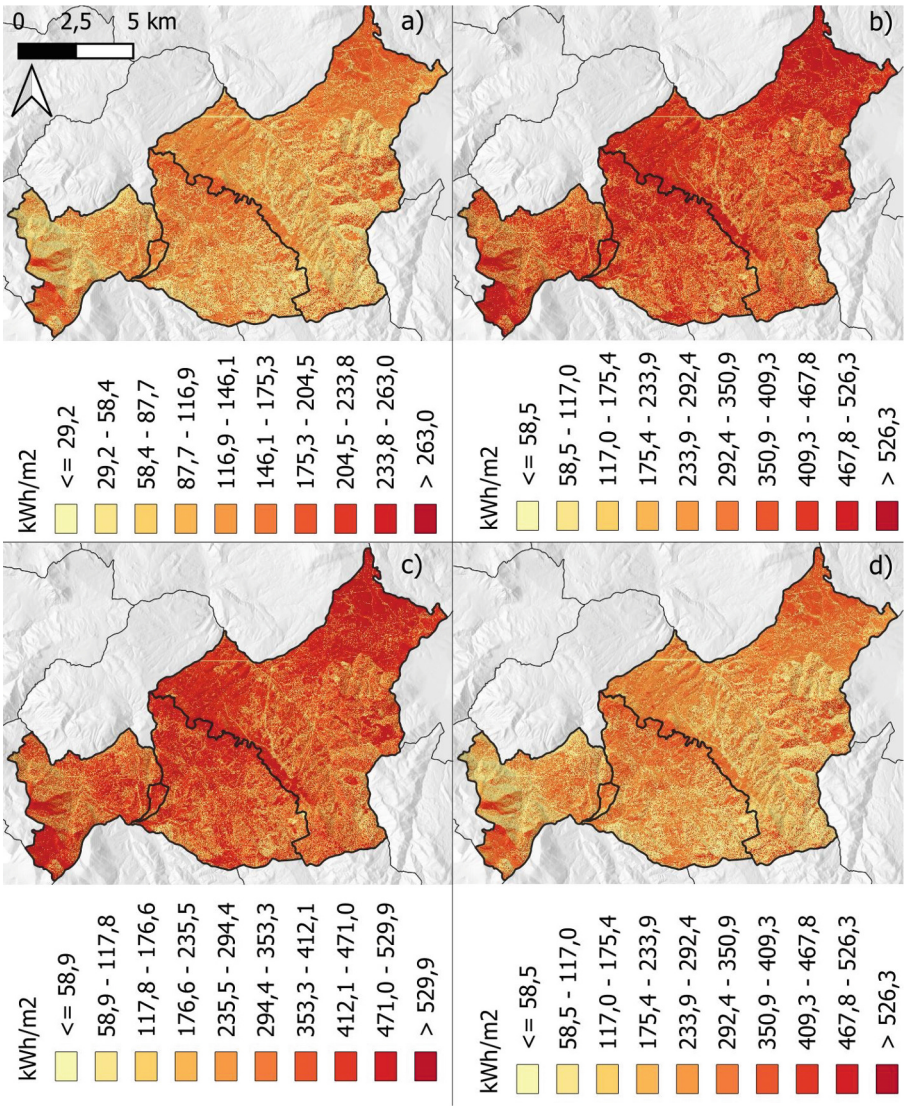


Fig. 5. Solar Radiation in (a) winter, (b) spring, (c) summer, (d) autumn.

4.4 Energy Production

In order to calculate the yield of the panels, at this point, each suitable building was associated with the average seasonal global solar radiation calculated as explained in Sect. 4.2 using zonal statistics. Then the energy producible (E_p) can be calculated as (Eq. 3)

$$E_p = R_{real} \cdot \eta \cdot PR \cdot \mathrm{A} \tag{3}$$

where

Table 3. Number of days with cloud (cirri, high and medium) probability clouds coverage percentage for each season.

Season	Cloud coverage %	Cloudy days	Season	Coverage %	Cloudy days
Winter	0–25%	28	Spring	0–25%	62
	25–50%	11		25–50%	9
	50–75%	24		50–75%	6
	75–100%	28		75–100%	14
Summer	0–25%	81	Autumn	0–25%	64
	25–50%	6		25–50%	8
	50–75%	2		50–75%	7
	75–100%	2		75–100%	12

Table 4. Data related to FV areas distribution in not-industrial or industrial built-up areas.

Coverage percentage	CRF	Meaning
0–25%	0.98	mostly clear skies, slight attenuation
25–50%	0.80	moderate attenuation
50–75%	0.55	heavy cloud cover
75–100%	0.25	very overcast sky

- η is the module efficiency, linked to the technology used in the panels and represents the ratio between the power output and the solar energy converted into electricity under STC conditions (ground irradiance of 1000 W/m^2, temperature of 25 °C and spectral distribution AM = 1.5).
- *PR* is a performance ratio of the panels. It provides information on the quality/yield ratio of the system, and therefore on system losses, such as those resulting from temperature, the inverter, mismatch situations, shading, soiling, etc. (related to system losses).
- *A* is the usable surface area (as described in Sect. 4.3) of suitable buildings (Table 5).

A summary of the technical literature in the field [15–17] shows that the average values of the indices mentioned can be those shown in Table 6. A conservative assumption was made when selecting these values in order to avoid overestimating the energy potential. Consequently, the actual energy production could be higher than the calculated values.

4.5 Energy Consumption

In order to estimate energy consumption, the average annual per capita electricity consumption was considered, equal to 5047kWh [18]. This one refers to total electricity

Table 5. Average η and PR values for Energy production.

Scenario	η	PR	Usage
Conservative	0.20	0.75	Regional studies with high uncertainty regarding shadows, dirt, BOS quality; or when you want to avoid overestimates
Central	0.22	0.80	Default for rooftops in Italy with modern modules and standard systems
Optimistic	0.24	0.85	Best-case scenario with premium modules (>24%) and excellent design/management (PR towards the high end of the modern range)

consumption and has been multiplied by the total number of inhabitants deduced from the ISTAT 2025 population data. The method used was intended as an estimate of electricity demand based on demographic proxies and was necessary due to the lack of disaggregated and updated data available at the sub-municipal level. This is consistent with the approach taken, which does not aim to provide a detailed analysis of sector energy loads but rather to summarise the overall electricity demand in order to compare it with energy production and understand whether the supply of PV, shifted to rooftops, can meet the demand. Table 6 shows the final calculated values, which identify a supply potential in terms of energy produced (E_p) that is significantly higher than the energy consumed (E_c).

Table 6. Conversion of the solar radiation in Energy producted (E_p), seasonally and annually and comparison with annual Energy consumption (E_c).

Season	Solar radiation (kWh/m^2)	E_p (kWh)	Annual E_p (kWh)	Annual E_c (kWh)
winter	874,761	13,768,300	152,794,600	53,361,931
spring	3,601,130	55,391,100		
summer	4,172,990	64,174,000		
autumn	1,237,340	19,461,200		

5 Discussion and Final Conclusion

This study aimed to propose an integrated approach that combines spatial analysis and remote sensing in order to verify if the installation of landscape-aware photovoltaic systems, which can be installed on building roofs, is capable of guaranteeing energy potential comparable to ground-mounted installations, thereby minimizing the consumption of agricultural land and the impact on the landscape. In light of the results obtained, it

can be said that this hypothesis is solidly supported by empirical evidence in the sample area analysed.

The analysis conducted with the aid of geospatial analysis tools (GIS) was based on the integration of seasonal solar radiation modelling, cloud cover correction using Sentinel-2 data processed in Google Earth Engine, and morphological and functional selection of building coverings, which made it possible to construct a realistic and prudent energy scenario. The application of CRFs to radiation in clear sky conditions further improved the reliability of the estimates, surpassing the simplified approach typical of some GIS modelling based on ideal atmospheric conditions.

The total area of suitable roofing identified (1,054,800 m^2), selected by excluding listed buildings belonging to the cultural heritage and located in historic centres, and buildings with morphological and dimensional limitations on roofing, as explained in Sect. 4.3, represents a "latent energy infrastructure" already integrated into the urban and productive fabric. The adoption of conservative parameters for module efficiency (η) and Performance Ratio (PR), consistent with the indications in the literature, has avoided overestimating the producible potential. It follows that the values obtained can be considered a minimum precautionary threshold.

The energy balance carried out by comparing producible energy (E_p) and consumed energy (E_c), this last estimated by using a demographic-proxy based on ISTAT population data for 2025, highlights a potential production that resulted significantly higher than the total electricity demand in the study area.

This result is particularly significant considering the context of European decarbonisation policies, outlined in Directive (EU) 2018/2001 and Regulation (EU) 2021/1119 on climate, which aim to achieve climate neutrality by 2050. In fact, this study shows that the objectives of increasing renewable energy sources can also be pursued through territorial strategies that do not involve further consumption of agricultural land or compromise the rural landscape. This responds to the critical issues frequently associated with large ground-mounted photovoltaic plant.

From a spatial planning perspective, this work aims to emphasise that buildings are no longer merely a passive element of urban metabolism, but are becoming an active support for energy transition. The available surface area of building roofs, when properly mapped and assessed, can form the backbone of local energy communities and distributed generation models, reducing territorial conflicts and promoting the social acceptability of interventions.

The methodology presented in this paper could also have further improvements. For example, there is the need to deepen the analysis by using disaggregated energy data, possibly at the building level. This would make possible to refine electric demand estimates, by also considering and integrating data related to socio-economic and performance variables of buildings and of people who live in them. Furthermore, the integration of photovoltaic production and energy efficiency in the building stock is an essential condition for maximising the environmental and social benefits of the transition.

Looking ahead, extending the analysis to other already impermeable surfaces (car parks, production areas, urban and peri-urban vacant lots) could further strengthen the

capacity to produce energy without taking away agricultural or natural land, consolidating an energy development strategy consistent with landscape protection and the principles of sustainable planning.

In conclusion, the results show that photovoltaic planning supported by GIS and geared towards enhancing existing buildings can offer energy potential comparable to ground-mounted solutions, significantly reducing the impact on the landscape and land consumption. From this perspective, the energy transition is not a process of spatial replacement, but rather an operation of reinterpretation and reactivation of existing territorial heritage, in line with a model of sustainability that integrates energy, landscape and local communities.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. EC: Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions "Clean Energy for all Europeans", COM (2016) 860 final. European Commission (2016)
2. EU: Directive 2018/2001 of the European Parliament and of the Council of 11 December 2018 on the promotion of the use of energy from renewable sources
3. EC: Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions the European "Green Deal", COM/2019/640 final. European Commission (2019)
4. Bevk, T., Golobič, M.: Contentious eye-catchers: perceptions of landscapes changed by solar power plants in Slovenia. Renew. Energy **152**, 999–1010 (2020)
5. Codemo, A., Ghislanzoni, M., Prados, M.-J., Albatici, R.: Incorporating public perception of renewable energy landscapes in local spatial planning tools: a case study in Mediterranean countries. Appl. Geogr. **170**, 103358 (2024)
6. Tamiminia, H., Salehi, B., Mahdianpari, M., Quackenbush, L., Adeli, S., Brisco, B.: Google Earth Engine for geo-big data applications: a meta-analysis and systematic review. ISPRS J. Photogramm. Remote Sens. **164**, 152–170 (2020)
7. Moore, R., Hansen, M.: Google Earth Engine: a new cloud-computing platform for global-scale earth observation data and analysis. In: AGU fall meeting abstracts, pp. IN43C-02
8. Ravanelli, R., et al.: Monitoring the impact of land cover change on surface urban heat island through Google Earth Engine: proposal of a global methodology, first applications and problems. Rem. Sens. **10**, 1488 (2018)
9. Mutanga, O., Kumar, L.: Google earth engine applications, vol. 11, p. 591. MDPI (2019)
10. Patel, N.N., et al.: Multitemporal settlement and population mapping from Landsat using Google Earth Engine. Int. J. Appl. Earth Obs. Geoinf. **35**, 199–208 (2015)
11. Huntington, J.L., et al.: Climate engine: cloud computing and visualization of climate and remote sensing data for advanced natural resource monitoring and process understanding. Bull. Am. Meteor. Soc. **98**, 2397–2410 (2017)
12. Fu, P., Rich, P.M.: A geometric solar radiation model with applications in agriculture and forestry. Comput. Electron. Agric. **37**, 25–35 (2002)
13. Hetrick, W.A., Rich, P.M., Barnes, F.J., Weiss, S.B.: GIS-based solar radiation flux models. In: ACSM ASPRS Annual Convention, American Soc Photogrammetry & Remote Sensing+ Amer Cong On p. 132

14. Fu, P.: A geometric Solar Radiation Model with Applications in Landscape Ecology. University of Kansas (2000)
15. IEA-PVPS: Uncertainty in Yield Assessments and PV Levelized Cost of Electricity (LCOE) International Energy Agency Photovoltaic Power Systems Programme (2020). https://iea-pvps.org/wp-content/uploads/2021/01/Report-IEA%E2%80%93PVPS-T13-18_2020-Uncertainties-in-Yield-Assessments-and-PV-LCOE-1.pdf
16. IEA-PVPS: Performance of Partially Shaded PV Generators Operated by Optimized Power Electronics. International Energy Agency Photovoltaic Power Systems Programme (2020). https://iea-pvps.org/wp-content/uploads/2021/01/Report-IEA%E2%80%93PVPS-T13-18_2020-Uncertainties-in-Yield-Assessments-and-PV-LCOE-1.pd
17. ISE, F.I.f.S.E.S.: Photovoltaics Report. Freiburg, Germany: Fraunhofer ISE (2024). https://www.ise.fraunhofer.de/content/dam/ise/en/documents/publications/studies/Photovoltaics-Report.pdf
18. IEA: International Energy Agency, IEA Energy Statistics Data Browser. International Energy Agency (2023)

Spatial Patterns of Urban Air Pollution Exposure in Cultural Heritage Areas Derived from Downscaled Satellite Observations

Maria Danese[1], Felice Perciante[2(✉)], and Rosa Lasaponara[2]

[1] ISPC-CNR, Cont. S. Loya, 85050 Tito Scalo, PZ, Italy
[2] IMAA-CNR; Cont. S. Loya, 85050 Tito Scalo, PZ, Italy
felice.perciante@cnr.it

Abstract. Satellite data on pollutants have too coarse a spatial resolution to be suitable for use in studying the exposure of cultural heritage. The integrated use of environmental and anthropogenic data with pollution data derived from Sentinel-5p using machine learning methods makes it possible to obtain high-resolution data that can be used to map urban pollution, which is useful for assessing the potential exposure of cultural heritage to atmospheric degradation.

Keywords: Sentinel-5p · Pollution · Cultural heritage · machine learning · spatial analysis · Rome

1 Introduction

Among the effects of the environmental emergency that characterises our times, one in particular is the cause of the acceleration of the material degradation of cultural heritage (CH): the increase in atmospheric pollution and in particular of gaseous compounds such as carbon monoxide (CO), sulphur dioxide (SO_2) and nitrogen dioxide (NO_2) [1]. This occurs mainly in urban contexts characterised by high anthropic pressure, resulting from high building density, high traffic volumes and the presence of large impermeable surfaces at the expense of green spaces [2, 3].

In general, these pollutants, even in combination with other environmental variables such as humidity, temperature and wind, lead to an increase in the phenomena of alteration [4] of stone materials and decorated surfaces, as well as acidification and corrosion of metals [5]. In the long term, this can compromise the physical integrity of cultural heritage.

More specifically, each of the aforementioned compounds acts differently on materials. For example, SO_2, when reacting with atmospheric humidity, generates sulphuric acid, which is particularly harmful to carbonate materials (marble, limestone). Within these materials, it causes an increase in porosity or even surface defects, which overall leads to an increase in the fragility of the material itself [6]. NO_2 acts in a similar way, causing acid rain when present in excess, as with SO_2, and the deposition of reactive compounds, as it acts as a catalyst for the oxidation of SO_2 and accelerates the formation

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 385–400, 2027.
https://doi.org/10.1007/978-3-032-30530-5_24

of crusts [7]. These react both on the surface materials of heritage sites, especially on natural stone, plaster and decorated surfaces [8]. CO is a less reactive pollutant, although, in the presence of other pollutants and other environmental parameters, including, once again, humidity, it contributes to an increase in degradation phenomena [9].

The distribution of CO, SO_2 and NO_2 values is strongly dependent on the spatial distribution of environmental parameters and anthropogenic variables, which can change even over relatively short distances. Therefore, effective mapping of these compounds on a scale appropriate to cultural heritage is very important for the assessment and prevention of cultural heritage degradation processes.

One of the primary sources useful for mapping compounds present in the air are ground-based monitoring networks such as the ARPA network in Italy [10]. However, these have the limitation of being point data, distributed in a limited and uneven spatial manner, despite attempts to make the data continuous through spatial interpolation methods [11, 12].

Another useful method for studying the distribution of atmospheric pollutants is the use of IoT networks, with distributed sensors [13]. However, these still have the limitation of being rather expensive in terms of infrastructure and management costs.

This is where Earth observation data from satellite missions come in, opening up new perspectives for atmospheric monitoring. More specifically, since 2018, as part of the Sentinel-5P TROPOMI mission of the Copernicus programme led by the European Commission in collaboration with the European Space Agency (ESA) [14], has been collecting data useful for assessing air quality. Although the data obtained from this mission has the advantage of offering global coverage and temporal continuity (with daily resolution) that other instruments do not have, it nevertheless has the limitation of having a spatial resolution, on the order of kilometres, which is not sufficient for applications on the scale of cultural heritage or, more generally, even on an urban scale. Furthermore, the data sometimes presents spatial gaps due to the presence of clouds or reflections from certain surfaces.

This work therefore proposes a method based on the use of machine learning techniques and the integration of heterogeneous and multi-source data for the reconstruction of data relating to the aforementioned pollutants, capable of reconstructing it in missing areas and increasing its spatial resolution thanks to its dependence on environmental, morphological and anthropogenic variables. In addition, an index will be introduced to assess the exposure of cultural heritage by evaluating its hot spots, thus creating a decision-making tool for monitoring and strategic thinking, as well as heritage protection. The case study is a complex urban context rich in cultural heritage, the Municipality of Rome, characterised by environmental conditions that are potentially critical for it.

2 Case Study

The area selected as the case study is the municipal territory of the city of Rome (Italy), which covers an area of approximately 1,285 km^2 and hosts over 2.8 million inhabitants. It represents an extremely complex and heterogeneous urban environment, characterized by the coexistence of a high density of cultural heritage assets—including globally significant sites such as the Colosseum, the Imperial Fora, and numerous basilicas—alongside a substantial building density and vehicular traffic. These elements alternate

with green spaces and industrial or commercial areas (Fig. 1). Within this context, urban morphology plays a key role in influencing the distribution of pollutants, thereby justifying the selection of the city as a case study.

In addition to urban morphology, the distribution of environmental parameters—and consequently of pollutants—is closely related to the city's geographical configuration, as it lies within the Tiber River basin and is surrounded by volcanic structures and a system of hilly reliefs.

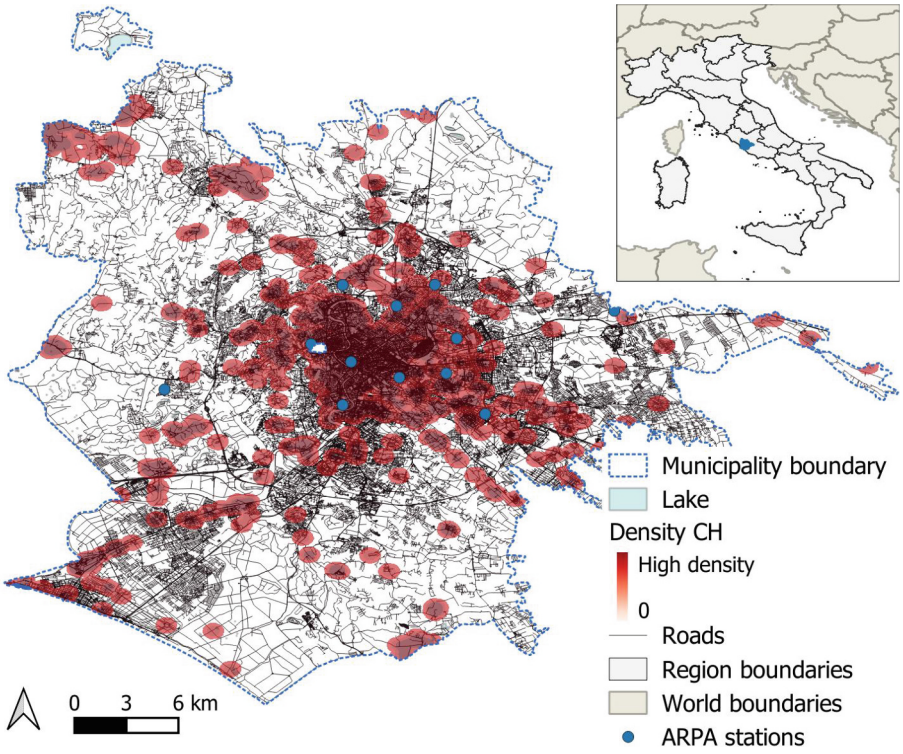


Fig. 1. The case study: the Rome Municipality with the Kernel Density Estimation of its cultural heritage. The first-order expected Nearest Neighbor was used as bandwidth in the estimation.

3 Materials and Methods

3.1 Input Datasets

In this study, the three S-5P datasets relating to the pollutants CO, NO_2 and SO_2 have a spatial resolution of 3.5 × 7 km and a daily temporal resolution, with a pass between 11:00 and 12:00 UTC (Coordinated Universal Time), corresponding to 13:00–14:00 local Italian time over the Rome area. These data consists of the instantaneous measurement of the vertical column density of the pollutant in the atmosphere, expressed in mol/m^2 [14].

As for the explanatory variables, both static and dynamic variables were included in the model. The static variables comprise groups of static morphological and anthropogenic parameters; the dynamic variables comprise environmental parameters.

With regard to the morphological parameters, the TINITALY DEM [15], was used, with an initial resolution of 10 m, which was downscaled to a resolution of 500 m using a mean focal function.

With regard to the anthropogenic data, the following variables were used:

- population and built-up density for 2025, taken from the GHSL (Global Human Settlement Layer) dataset of the Copernicus Emergency Management Service, developed by the European Commission's Joint Research Centre, with a spatial resolution of 100 m, to map the population and human settlements.
- road and industrial density in 2025. Roads and industrial sites were derived from OpenStreetMap (OSM) [16] via the website https://overpass-turbo.eu/ (tags: roads, industry), and the densities were subsequently calculated as road length in km per pixel and as the number of industrial buildings per pixel, respectively.

For what concerns the environmental parameters, the following have been included in the model:

- climatic data derived from ERA5-Land Daily, developed by the European Centre for Medium-Range Weather Forecasts (ECMWF) as part of the Copernicus Climate Change Service (C3S) programme [17]. More specifically, the following variables were considered: average, minimum and maximum air temperature at 2 m, surface pressure (measured in Pa), relative humidity (%), wind speed at 10 m (in m/s), and average solar radiation at ground level (W/m^2). These data have a spatial resolution of 9 km and a daily update frequency.
- daily cumulative precipitation derived from CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) [18], with a spatial resolution of 5 km.
- Dynamic World V1 land cover, derived from Sentinel products by the World Resources Institute (WRI) [19], with a spatial resolution of 10 m and daily temporal resolution, containing the following land cover classes: water, trees, grass, flooded vegetation, crops, shrubs and scrub, built-up areas, bare ground, snow and ice.
- NDVI index, calculated as the normalised combination of the red (RED) and near-infrared (NIR) bands from the Sentinel-2 (S2) dataset, it belongs to the mission of the European Copernicus programme [20]. It ranges from 1 to -1, and its values indicate changes in the health and density of vegetation. It has a temporal resolution of approximately 5 days and a spatial resolution of 10 m.

The model utilised environmental parameters from the day prior to the passage of S-5P, as the formation and changes in pollutant data are not instantaneous but depend on the weather conditions (temperature, rainfall, wind) of the preceding hours.

Finally, the data on Rome's CH were also derived from OSM.

3.2 Processing Environment and Modelling Approach

Part of the data processing and modelling workflow mentioned in the previous section was carried out within Google Earth Engine (GEE). This is a cloud-based platform for

processing and analyzing large-scale geospatial data, which has the following advantages: high-speed parallel processing without the need for large computing resources on one's own computer, easy access to data and methods for different types of users [21, 22]. It is therefore particularly suitable for processing satellite images and analyzing climate datasets [23, 24]. In fact, it provides free access to a vast dataset [25], including all the data mentioned in the previous section.

Within GEE there are several algorithms that can be used to analyze spatial and satellite data, and among these are some machine learning methods such as supervised methods like the support vector machine or the random forest (RF) [26, 27], used in this work for the prediction of missing values (nodata) in air pollutant data derived from the Sentinel-5P satellite.

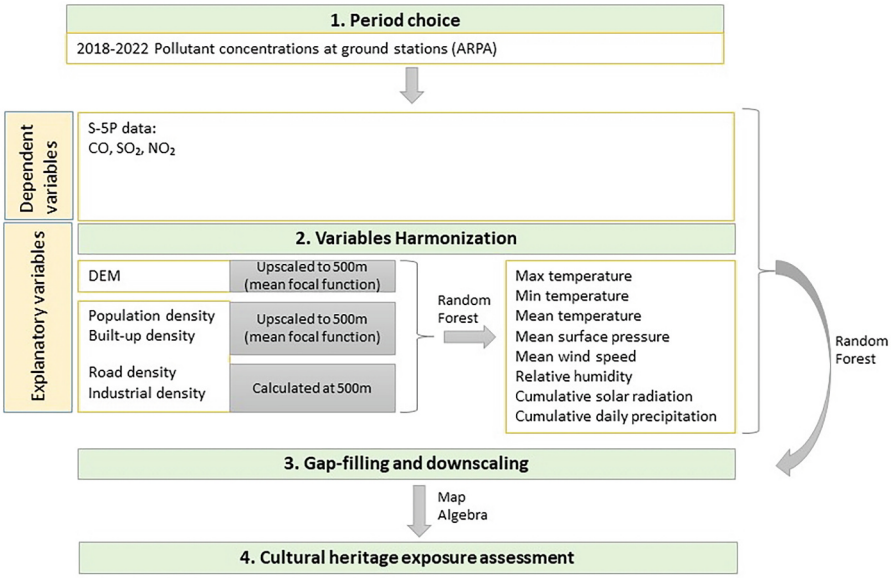


Fig. 2. Flow chart resuming the process.

The RF method is an ensemble-type supervised learning algorithm, based on the optimization of a set of decision trees following recursive execution of the algorithm several times; this improves the predictive ability of the model by reducing data overfitting [28, 29]. A training set must be chosen within the population dataset, then the decision tree classification is performed. Finally, as anticipated, the process is repeated to improve the learning process and return the best classification.

Instead, the calculation of statistics on gap-filled and downscaled data and the estimation of exposure were carried out using map algebra methods [30], a set of mathematical and logical operations applied to raster data, capable of performing calculations on a cell-by-cell basis (local functions), within predefined zones (zonal functions) or within the neighborhoods of each pixel (focal functions).

Figure 2 shows a flowchart summarising the data and the method followed, the steps of which will be explained in the following section on the results.

4 Results

4.1 Period Choice

For the aim of the analysis presented in this study, the first step involved examining the temporal trends in pollutant concentrations, derived from data collected during the ARPA Lazio air quality monitoring campaigns between 2018 and 2025. In the Rome area, specifically, there are 11 stations, whose pollutant data are recorded hourly, daily and monthly. Figure 3 illustrates the trends for each station measuring the concentrations of each pollutant under consideration; the figure reveals the presence of seasonal patterns. More specifically, the maximum peaks for NO_2 and CO occur during the winter months. This may be linked to the combined conditions arising from the increase in emissions during this period due to heating and road traffic, as well as meteorological conditions that favour the stagnation of pollutants in the lower atmospheric layers. In contrast, higher concentrations of SO_2 are observed in summer.

Consequently, this study focused on the month of January for NO_2 and CO data, whilst July was selected for SO_2 data, with the aim of analyzing conditions that are more representative of high environmental pressure, in order to identify the most significant spatial patterns of the pollutants and, consequently, the areas where the heritage is actually exposed.

4.2 Variable Harmonization

Once the study period has been selected, it is necessary to spatially harmonize the explanatory variables in order to ensure the spatial consistency of the predictor stack. The aim is to standardize all predictive variables to a spatial resolution of 500m, a value suitable for urban-scale analyses of CH, in line with the objectives of this study.

Initially, all static information layers with high spatial resolution, along with the NDVI – namely the DEM, the NDVI, population density and built-up density – were upscaled using a mean focal function from map algebra. The exception is the land cover derived from Dynamic World, as it is a categorical variable, for which a majority focal function was used.

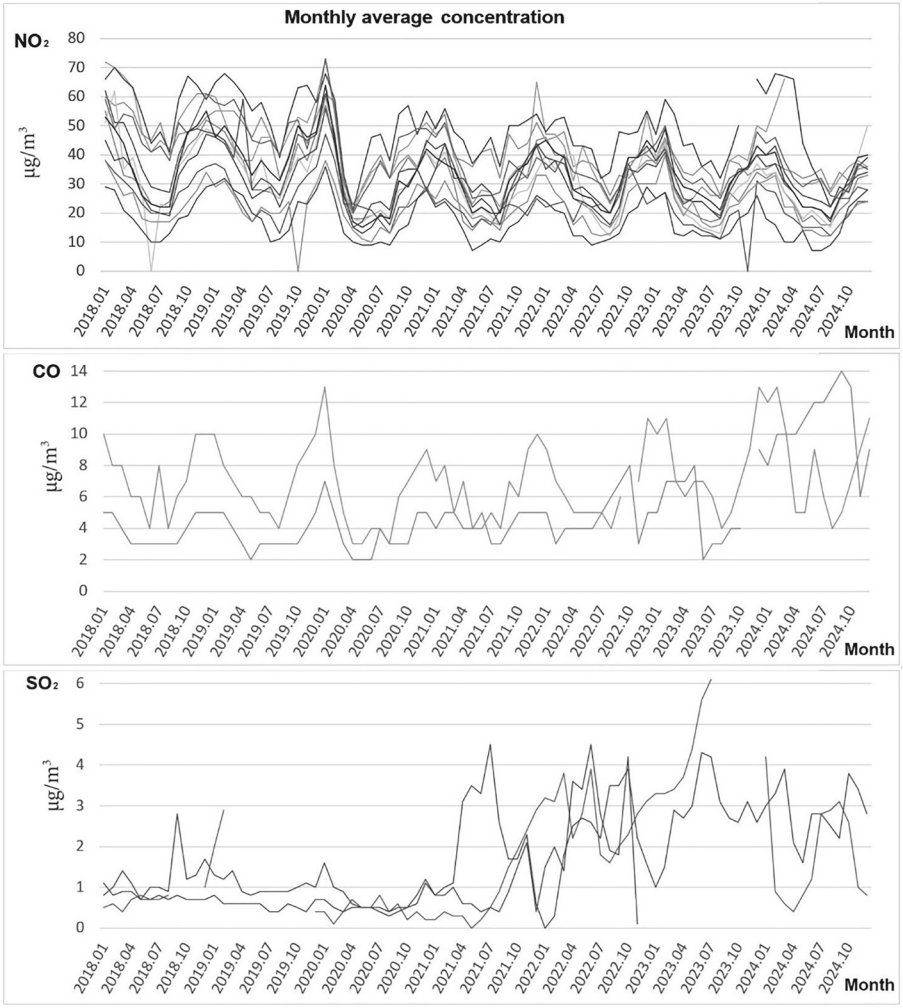


Fig. 3. Monthly pollution trends. Each line represents a meteorological station. Overall, the seasonal trends are evident.

These elements constitute the morphological, environmental and anthropogenic bases that characterize the territory and influence meteorological conditions; therefore, they were used as input datasets for the downscaling of environmental parameters derived from ERA5 and CHIRPS data. The method used was RF, with 100 trees. The result is a dataset of variables with a resolution of 500 m which nevertheless retains the information content of the input parameters, whilst improving the representation of local gradients (Fig. 4).

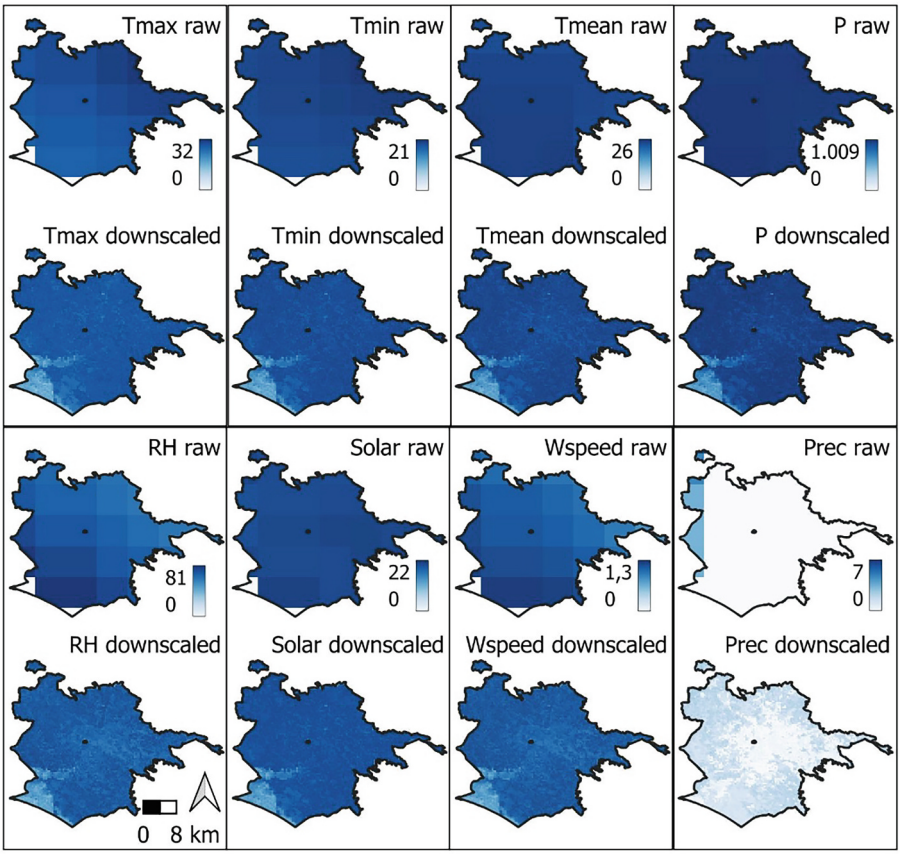


Fig. 4. Example of environmental parameters raw and downscaled.

4.3 Pollution Gap-Filling and Downscaling

Once the dataset had been prepared with the dependent and independent variables, the missing values in the atmospheric column for the three selected pollutants were imputed using the Random Forest algorithm within GEE, thereby producing a continuous map. Furthermore, this process improved the spatial resolution of the pollution data to 500 m (Fig. 5).

The results were validated by calculating the root mean square error (RMSE) and Pearson's correlation coefficient (r). The low values of the former and high values of the latter (Table 1) indicate that the model has good predictive capability, ensuring that the downscaled data remain consistent with the original satellite data. This confirms the robustness of the method.

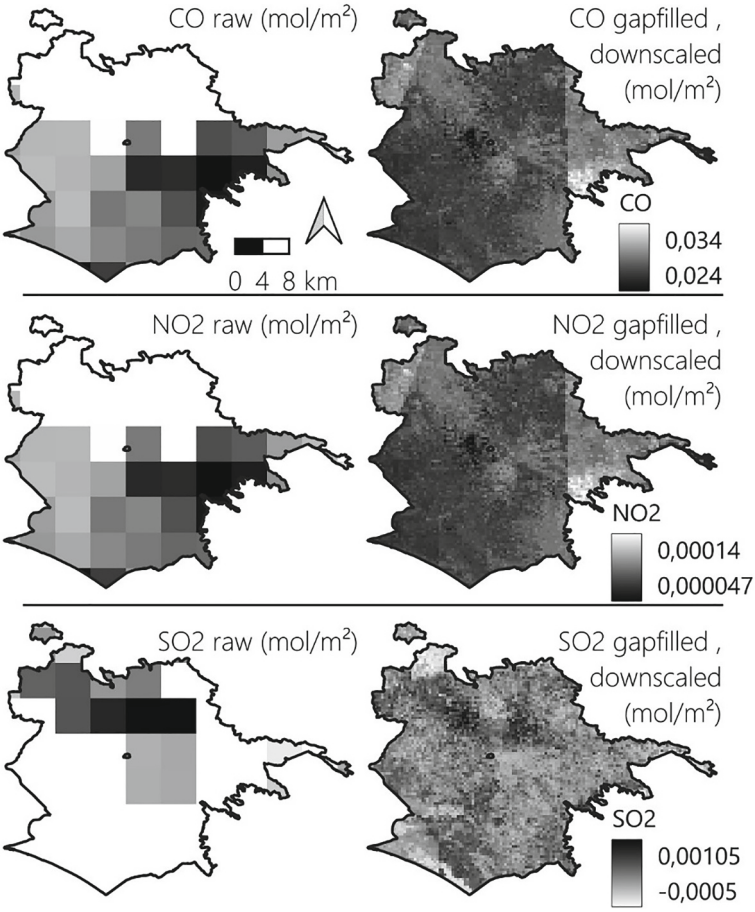


Fig. 5. Pollutants raw and gap filled-downscaled rasters.

Table 1. Range found for RMSE and Pearson's correlation (r).

Pollutant	RMSE range (mol/m^2)	r range
CO	0.0001–0.0014	0.86–0.96
NO_2	$10^{-6} - 10^{-5}$	0.85–0.95
SO_2	0.0001– 0.00045	0.90–0.94

4.4 Cultural Heritage Exposure Assessment

Once the continuous, high-resolution rasters of the pollutants had been obtained, a raster of average values was calculated for each pollutant, using the local statistics of map

algebra, for each pixel within the month in question. The same was done with the maximum values. The first result provides information on the background concentration of the pollutant, as it describes an average monthly level of pollution. It can therefore be associated with slow but cumulative and long-term degradation processes affecting CH. The second result, on the other hand, provides information on peaks occurring during the period and thus on locations of episodic but intense exposure. It therefore implies phenomena of accelerated degradation processes, indicating areas with a higher risk, albeit for shorter and rarer periods. The combined use of the two indicators thus helps to provide a more complete picture of the heritage's exposure.

Subsequently, for the purpose of calculating a CH exposure index, a Euclidean distance was calculated from the CH sites, which was then normalised.

The following Heritage Exposure Index (HEI) was calculated from the combination of the normalised pollutants and the normalised distance from the heritage sites, Eq. 1:

$$HEI = pollutant\ concentration\ x\ (1 - proximity\ to\ heritage\ sites) \quad (1)$$

The result of this equation returns: (1) a value approaching 1 (worst-case scenario), if there is a high concentration in the vicinity of assets; (2) a value approaching 0 (best-case scenario), if there is a low concentration at a distance from assets; (3) intermediate values that occur in the case of low concentration near assets or in the case of high concentration at a distance.

The HEI values were classified into five categories using the quantile method (Figs. 6, 7 and 8), in order to highlight the relative distribution of exposure within the study area and identify the areas characterised by higher levels of severity.

Analysis of the findings from the HEI index shows that the pollutants analysed exhibit different spatial patterns.

As regards CO, the concentrations in the 'Very high' and 'High' classes of average values are found in the central area of Rome, which accounts for 60% and 33% of the CH respectively. This is undoubtedly due to the widespread presence of these pollutants in the city's most densely populated urban areas. Incidents of maximum CO levels are also found predominantly in areas classified as 'High'.

NO_2 levels show a similar pattern, with different percentage values but a comparable trend, again linked to road traffic and urban density.

SO_2 behaves differently, however, both in terms of average values – which are concentrated mainly in the Medium category – and maximum values, which are found predominantly in the Low and Very Low categories. This is likely due to the spatial dispersion of emission sources and an overall lower concentration of the pollutant (Table 2).

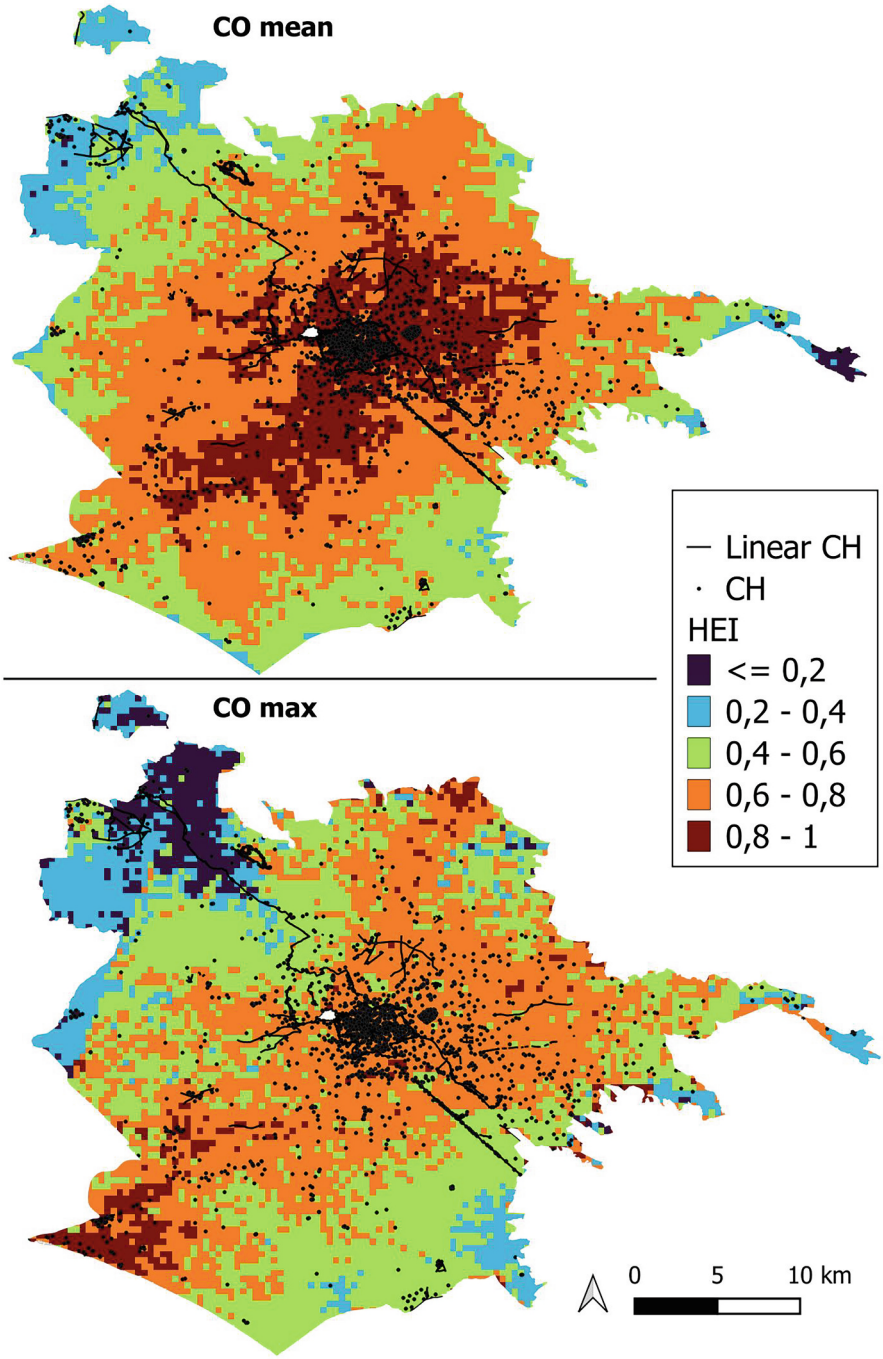


Fig. 6. HEI index calculated for CO mean and maximum.

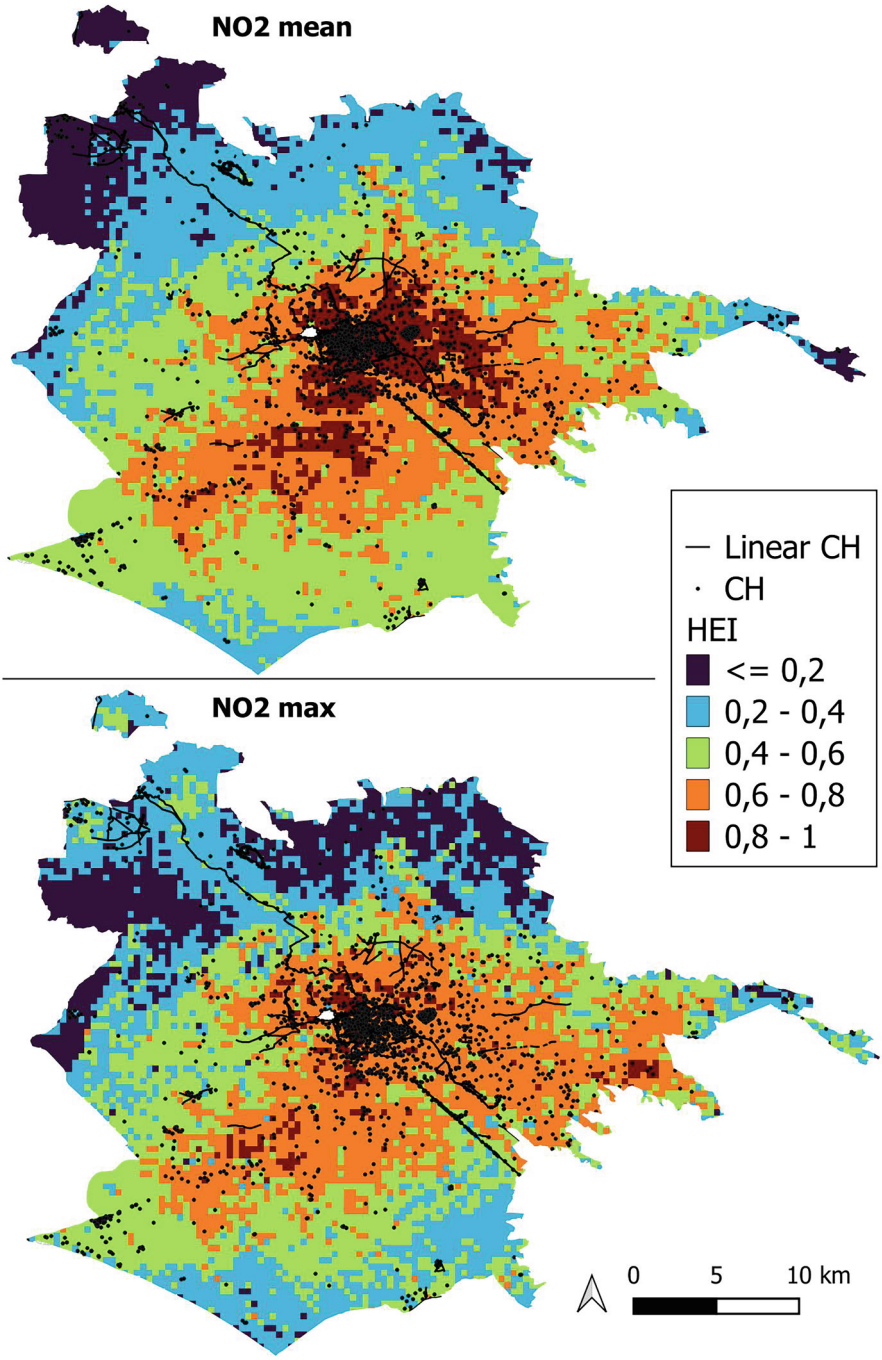


Fig. 7. HEI index calculated for NO_2 mean and maximum.

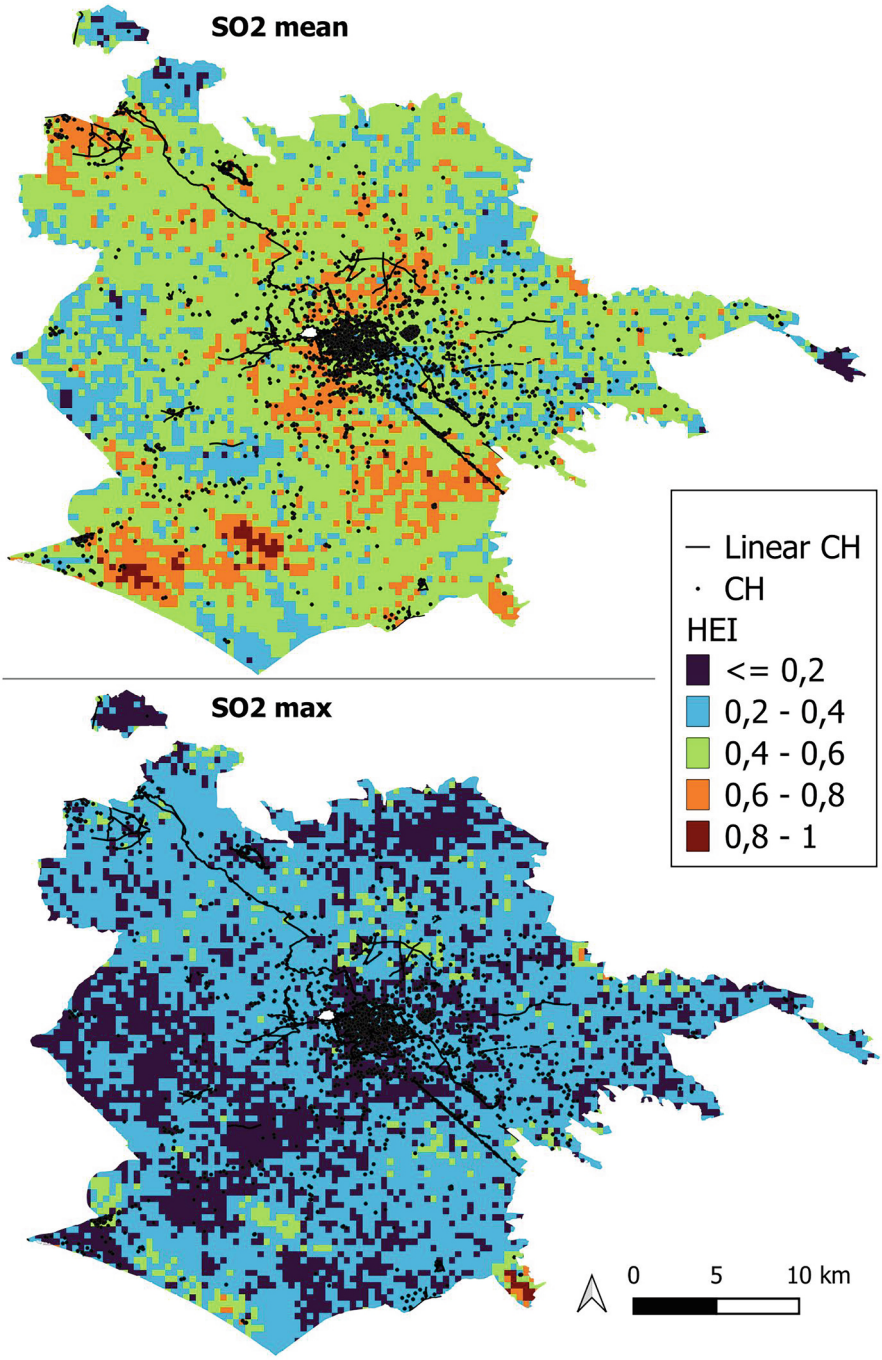


Fig. 8. HEI index calculated for SO_2 mean and maximum.

Table 2. CH percentage (%) following in HEI classes for each pollutant.

HEI	CO mean	CO max	NO_2 mean	NO_2 max	SO_2 mean	SO_2 max
Very low	0.03	0.46	0.68	2.2	0.11	40.4
Low	0.36	2.66	4.8	7.37	12.93	54.44
Medium	5.79	19.47	16.09	18.05	72.19	3.40
High	33.42	74.59	29.30	61.99	14.69	0.72
Very high	60.4	2.81	49.14	10.39	0.07	1.03

5 Discussion and Conclusions

This study presents an approach aimed at assessing the exposure of the CH to some of the main atmospheric pollutants, based on the integration of remote sensing, machine learning and spatial analysis. Overall, the approach adopted has demonstrated how the integration of technologies and methods can help to improve the available data and carry out a more effective analysis at an urban scale.

The introduction of the HEI index has made it possible to combine information on pollutants with the location of properties, providing a representation and quantification of exposure levels. This analysis revealed that, in the context under investigation, the main exposure factors are attributable to CO and NO_2, whilst the contribution of SO_2 is more limited. The indicator provides a useful decision-support tool for programming and planning, helping to identify priority areas for monitoring and preventive conservation.

The method could be further developed by incorporating the contribution of other pollutants, extending the analysis to longer time series, and including the specific vulnerability of different materials in the analysis.

However, the proposed method constitutes an easily replicable and scalable protocol to support the protection of cultural heritage in a context of increasing environmental pressure.

Acknowledgments. IR0000032 – ITINERIS, Italian Integrated Environmental Research Infrastructures System (D.D. n. 130/2022 - CUP B53C22002150006) Funded by EU - Next Generation EU PNRR- Mission 4 “Education and Research” - Component 2: “From research to business” - Investment 3.1: “Fund for the realization of an integrated system of research and innovation infrastructures”. Project “Study of hazard elements and their variation as a function of climate change, on CH using spatial ML methods” inside “CH and creativity in green and digital transitions for inclusive societies” (FOE2021 - Ordinary fund for research organizations and institutions).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Tidblad, J., et al.: Effects of air pollution on materials and cultural heritage: ICP materials celebrates 25 years of research. Int. J. Corrosion **2012**, 496321 (2012)

2. Vardoulakis, S., Fisher, B.E., Pericleous, K., Gonzalez-Flesca, N.: Modelling air quality in street canyons: a review. Atmos. Environ. **37**, 155–182 (2003)
3. Nowak, D.J., Hirabayashi, S., Bodine, A., Greenfield, E.: Tree and forest effects on air quality and human health in the United States. Environ. Pollut. **193**, 119–129 (2014)
4. Esteban-Cantillo, O.J., Menendez, B., Quesada, B.: Climate change and air pollution impacts on cultural heritage building materials in Europe and Mexico. Sci. Total. Environ. **921**, 170945 (2024)
5. Brimblecombe, P., Grossi, C.M.: Damage to buildings from future climate and pollution. APT Bull. J. Preserv. Technol. **38**, 13–18 (2007)
6. Charola, A.E.: Acid rain effects on stone monuments. J. Chem. Educ. **64**, 436 (1987)
7. Bonazza, A., Messina, P., Sabbioni, C., Grossi, C.M., Brimblecombe, P.: Mapping the impact of climate change on surface recession of carbonate buildings in Europe. Sci. Total. Environ. **407**, 2039–2050 (2009)
8. Ausset, P., et al.: Experimental study of limestone and sandstone sulphation in polluted realistic conditions: the Lausanne Atmospheric Simulation Chamber (LASC). Atmos. Environ. **30**, 3197–3207 (1996)
9. Graedel, T.E., Hawkins, D.T., Claxton, L.D.: Atmospheric Chemical Compounds: Sources, Occurrence and Bioassay. Elsevier (2012)
10. Zavatti, A., Lauriola, P.: The contribution of environmental monitoring in the epidemiological assessment of exogenous risk: the experience of ARPA in the Emilia-Romagna Region of Italy. Ann. N. Y. Acad. Sci. **895**, 173–187 (1999)
11. Bechle, M.J., Millet, D.B., Marshall, J.D.: Remote sensing of exposure to NO2: satellite versus ground-based measurement in a large urban area. Atmos. Environ. **69**, 345–353 (2013)
12. Jerrett, M., et al.: A review and evaluation of intraurban air pollution exposure models. J. Eposure Sci. Environ. Epidemiol. **15**, 185–204 (2005)
13. Channi, H.K., Kumar, R.: The role of smart sensors in smart city. In: Singh, U., Abraham, A., Kaklauskas, A., Hong, T.P. (eds.) Smart Sensor Networks. Studies in Big Data, vol. 92. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-77214-7_2
14. Veefkind, J.P., et al.: TROPOMI on the ESA Sentinel-5 Precursor: a GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. Remote Sens. Environ. **120**, 70–83 (2012)
15. Tarquini, S., et al.: TINITALY/01: a new triangular irregular network of Italy. Ann. Geophys. **50**, 407–425 (2007)
16. Nelson, A., de Sherbinin, A., Pozzi, F.: Towards development of a high quality public domain global roads database. Data Sci. J. **5**, 223–265 (2006)
17. Hersbach, H., et al.: The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. **146**, 1999–2049 (2020)
18. Funk, C., et al.: The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci. Data **2**, 1–21 (2015)
19. Brown, C.F., et al.: Dynamic World, Near real-time global 10 m land use land cover mapping. Scientific Data **9**, 251 (2022)
20. Drusch, M., et al.: Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sens. Environ. **120**, 25–36 (2012)
21. Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R.: Google earth engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. **202**, 18–27 (2017)
22. Sangeetham, R.: a review on google earth engine: an open access cloud analysis platform for planetary scale satellite data. In: E3S Web of Conferences, p. 09011. EDP Sciences
23. Abijith, D., et al.: Assessing the impact of climate and land use change on flood vulnerability: a machine learning approach in coastal region of Tamil Nadu India. Geosci. Lett. **12**, 1 (2025)

24. Khalil, M., Kumar, J.S.: Predictive modeling of land surface temperature dynamics in Damascus, Syria using google earth engine: a remote sensing and random forest approach. Earth Sci. Inf. **18**, 1–20 (2025)
25. Tamiminia, H., Salehi, B., Mahdianpari, M., Quackenbush, L., Adeli, S., Brisco, B.: Google Earth Engine for geo-big data applications: a meta-analysis and systematic review. ISPRS J. Photogramm. Remote Sens. **164**, 152–170 (2020)
26. Ouchra, H., Belangour, A., Erraissi, A.: Exploring google earth engine platform for satellite image classification using machine learning algorithms. In: Ben Ahmed, M., Boudhir, A.A., El Meouche, R., Karaş, İ.R. (eds.) Innovations in Smart Cities Applications Volume 7. SCA 2023. Lecture Notes in Networks and Systems, vol. 938. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-54376-0_24
27. Zhao, Z., et al.: Comparison of three machine learning algorithms using google earth engine for land use land cover classification. Rangel. Ecol. Manage. **92**, 129–137 (2024)
28. Kleinberg, E.M.: On the algorithmic implementation of stochastic discrimination. IEEE Trans. Pattern Anal. Mach. Intell. **22**, 473–490 (2002)
29. Kleinberg, E.: An overtraining-resistant stochastic modeling method for pattern recognition. Ann. Stat. **24**, 2319–2349 (1996)
30. Tomlin, C.D.: Geographic Information Systems and Cartographic Modeling. Prentice Hall Englewood Cliffs, NJ (1990)

Scientific Computing Infrastructure (SCI 2026)

Optimizing Harmful Information Containment in Social Networks via Shortest-Path Interdiction

Muon Ha[1(✉)], Yulia Shichkina[2], Tran Duc Manh[2], Xuan -Hien Nguyen[3], and Thanh-Son Phan[4]

[1] Telecommunications University, Khanh Hoa, Vietnam
muon.ha@mail.ru
[2] St. Petersburg State Electrotechnical University, St. Petersburg, Russia
[3] Hanoi University of Industry, Ha Noi, Vietnam
[4] Thai Binh Duong University, Nha Trang, Khanh Hoa, Vietnam

Abstract. The spread of malicious information on social networks is often explosive and difficult to control, especially in small-world structures and echo chambers. Traditional intervention strategies based on degree centrality are effective in scale-free networks; however, they often perform poorly in relatively homogeneous community networks, where bridge nodes may be more important than hubs. In this paper, we propose SPID (Shortest-Path Interdiction for Diffusion), a corridor-based mitigation framework that targets fast-spreading "diffusion corridors", defined as shortest paths connecting high-risk source regions to target communities. The experimental implementation uses SPID++, a residual-updated version of SPID that adapts to rerouted paths after each intervention step. Experiments on the Watts–Strogatz network show that SPID++ reduces the total number of infected nodes by approximately 2.2 times compared with Degree Centrality. Budget sensitivity analysis further shows that SPID++ can protect up to 444 more nodes than Degree as the intervention budget increases. Runtime results indicate that batched shortest-path computation reduces the intervention-selection time of SPID++ while preserving the selected intervention set.

Keywords: shortest path · shortest-path interdiction · harmful information diffusion · small-world networks · intervention optimization · SEIR model

1 Introduction

The development of social media has created an ideal environment for the rapid spread of information. Alongside its positive benefits, social media has also become fertile ground for harmful content, including fake news, online scams, and the manipulation of public opinion. The core danger of this type of information lies in its "chain reaction" spread mechanism: from a small group of initiating nodes (seeds), information can quickly spread through social media links, creating infection peaks and exceeding the control capabilities of manual censorship systems.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 403–414, 2027.
https://doi.org/10.1007/978-3-032-30530-5_25

In cybersecurity management, the core objective is not only simply to remove content, but also to optimize prevention strategies to simultaneously reduce infection peaks and minimize the total number of affected individuals with the lowest intervention costs. Current common approaches often rely on centrality metrics to identify critical nodes and prioritize interventions. In this context, a Degree Centrality-based strategy (intervening with KOLs/hubs) is often considered effective for scale-free networks (like Twitter), where a few high-level nodes can strongly influence the flow of information. However, reality shows that social network structures are not always hub-centric. In closed groups, professional forums, or local communities, the network structure is often small-world, where dissemination relies more on "shortcuts" connecting clusters than on a few prominent individuals. In such cases, Degree Centrality becomes less effective because it is difficult to correctly identify the nodes acting as "bridges" and the crucial "corridors" for dissemination (Fig. 1).

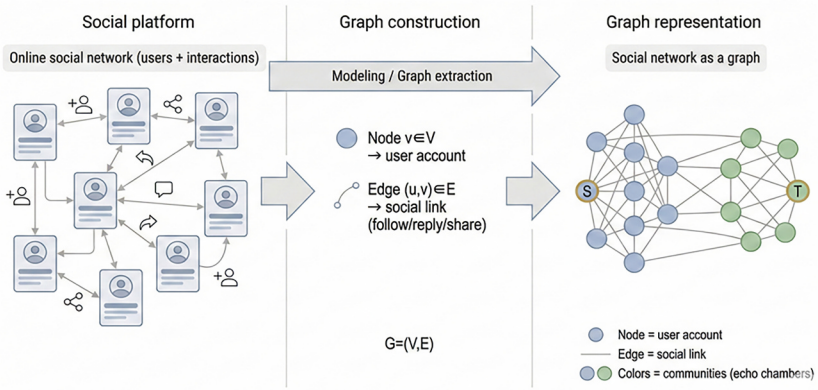


Fig. 1. Graph-based modeling of social media diffusion

A theoretically suitable approach to identify bridge nodes is Betweenness Centrality. However, this method faces significant computational obstacles when applied to large-scale graphs, making it impractical for networks with millions of nodes.

To fill this gap, the paper proposes a new method called SPID (Shortest-Path Interdiction for Diffusion). Instead of focusing on high-order nodes, SPID models the problem from a "path-blocking" perspective: finding and eliminating nodes on the shortest paths connecting the area of misinformation dissemination with vulnerable communities. This approach aims to cut off the fastest propagation routes, thereby reducing the spread of malicious information within the network.

The main contributions of the paper include: (i) proposing an algorithmic framework for node selection based on risk-aware path sampling and residual update mechanisms to adapt to changes in the information flow after intervention; (ii) using parallel shortest-path algorithms on large graphs as the core computation to address the performance issues encountered by Betweenness; and (iii) evaluating SPID on clustered network structures (Watts–Strogatz) and analyze budget-based sensitivity, thereby clarifying its potential application in scenarios where traditional methods perform poorly.

The paper is organized as follows. A review of related work and existing methods is presented in Sect. 2. Section 3 details the proposed parallel algorithm and its technical refinements. Experimental results obtained on large-scale datasets are reported and analyzed in Sect. 4. Finally, Sect. 5 concludes the paper and discusses possible avenues for future research.

2 Related Works

Recent research indicates that the rapid spread of misinformation is largely determined by the topology of the underlying network, such as community structures, echo chambers, and bridging nodes. These findings suggest that network-based models provide more accurate representations than traditional well-mixed population assumptions. Empirical data further demonstrate that fake news often diffuses faster and reaches a wider audience than factual information [1, 2]. Additionally, social bots act as significant amplifiers of untrustworthy content, particularly when such information aligns with existing community biases or current trends [3]. Furthermore, the "complex contagion" theory emphasizes the role of repeated exposure in clustered neighborhoods, increasing the probability of acceptance/sharing and contributing to widely spread dynamics [4].

To describe behavioral latency, SEIR-type models are often used because the Exposed state represents the "received but not shared" situation. This choice is suitable for modern studies that combine propagation with social context and temporal evolution. In terms of data, FakeNewsNet integrates content, social context, and spatiotemporal signals, supporting simultaneous analysis of propagation, detection, and intervention [7]. In terms of detection, surveys have systematized feature groups (content, users, propagation) and method families (classical machine learning, deep learning, multimodal learning, and explainable models) [5, 6]. For rumors, specialized reviews highlight the lack of early-stage data, label uncertainty, and the importance of conversational signals and community responses [8]. More recently, detection systems are able to interpret and exploit user content and feedback to provide "evidence" for decision-making have also been proposed [5, 6].

Mitigation methods can be broadly divided into "soft" and "hard" interventions. Soft interventions – for example, "accuracy nudges" – improve the quality of sharing without directly blocking content [9], while prebunking/inoculation increases resistance to manipulation techniques, even in the context of COVID-19 misinformation [10]. Community-based strategies also apply crowdsourcing to integrate detection and mitigation through collective labeling and feedback [11]. In contrast, hard interventions often model the intervention site selection problem as a network optimization problem under budget constraints, aiming to reduce exposure or spread intensity. Typical approaches involve point process-based intervention models [13], propagation control under diffusion processes [12], and recent extensions incorporating fairness/social constraints [14], time-based penalties under heterogeneous propagation rates [15], or adaptive selection of correctors through learning-to-intervene methods [16].

From a network optimization perspective, the network interdiction problem provides a bilevel framework: the upper layer selects nodes/edges to break or degrade under budget constraints, and the lower layer optimizes a network goal on the post-intervention graph

(e.g., shortest path or maximum flow). A recent survey has systematized interdiction models and algorithms, in which the shortest-path interdiction is a classic and widely studied case [17]. During the period 2016–2025, shortest-path interdiction has been expanded in directions such as arc improvements and structural analysis [18], settings with uncertainty and information asymmetry [19], incomplete information regimes [20], and policy learning approaches to increase scalability [21]. Notably, misinformation control has been modeled as a two-tiered stochastic interdiction problem to minimize expected reach to vulnerable groups [22], thereby reinforcing the link between breaking structural "corridors" and limiting propagation.

Finally, shortest-path/interdiction-based strategies often require iterative shortest-path computations across multiple post-intervention graph instances that make computational efficiency a key factor. Recent acceleration efforts leverage graph-algebra frameworks (e.g., GraphBLAS) to implement efficient SSSP on parallel hardware (e.g., delta-stepping) [23]. Complementary to this direction, an SSSP method tailored to graphs with cycles is introduced in [24] and subsequently extended to parallel execution on large graphs in [25], providing a practical implementation basis for shortest-path-based interventions.

3 Methodology

3.1 Problem Statement

Consider a directed graph $G = (V, E)$, where V is the set of nodes (e.g., user accounts) and E is the set of edges (e.g., interactions). Let $S_0 \subseteq V$ denote the initial set of sources that disseminate harmful information. Diffusion follows an SEIR (Susceptible–Exposed–Infected–Recovered) process with parameters $\Theta = \{\beta, \kappa, \gamma\}$, representing the exposure, infection, and recovery rates, respectively.

An intervention strategy selects a set I of elements to disable subject to a budget constraint: the total intervention cost over I does not exceed budget B, where each element x (node or edge) has cost $c(x)$. We consider two intervention modes:

- **Node interdiction**: $I \subseteq V$ (e.g., suspending accounts to reduce dissemination capacity).
- **Edge interdiction:** $I \subseteq E$ (e.g., weakening or removing transmission interactions).

After intervention, the network becomes $G' = (V', E')$. Motivated by shortest-path interdiction [17, 18, 22], effective interventions disrupt fast "diffusion corridors" by increasing distances from sources to target communities or, equivalently, reducing reachability.

We formulate the problem with one of two primary objectives:

Structural objective: maximize the aggregate shortest-path distance from sources S_0 to a target set T_0 (optionally via a monotone distance-based utility, with weights if needed).

Dynamical objective: minimize the number of infected nodes at equilibrium under SEIR simulation on G'.

Recent studies connect these views through interdiction-based diffusion control, modeling reachability or dynamics under budget constraints [17, 18, 22]. Section 3 details how shortest-path structure can be exploited to reduce SEIR spread.

3.2 The Proposed SPID Method

SPID (Shortest-Path Interdiction for Diffusion) selects an intervention set I (nodes or edges) under a finite budget B to reduce diffusion in SEIR simulations. Unlike approaches that prioritize globally central nodes via static centrality measures, SPID targets "diffusion corridors" approximated by shortest paths from sources S_0 to targets T_0. Figure 2 illustrates the simulated social network and SPID's core intuition: interdiction of shortest-path corridors.

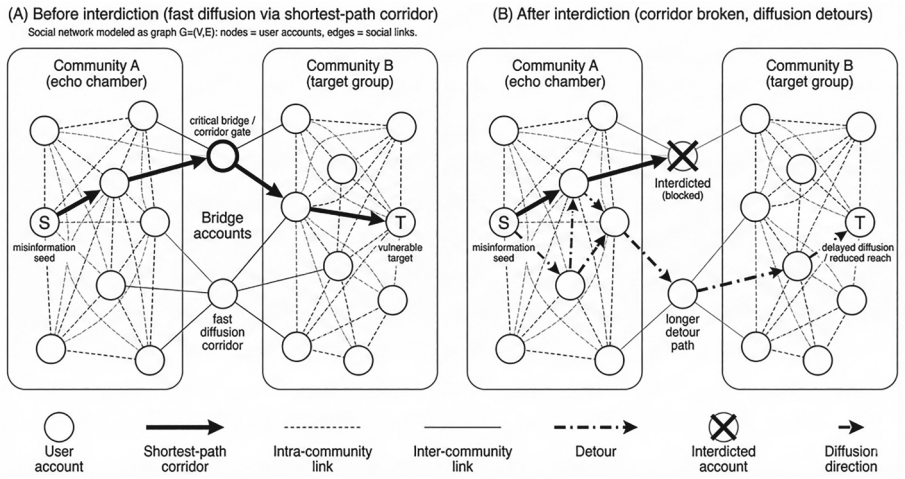


Fig. 2. Social network graph simulation and SPID intuition: shortest-path corridor interdiction

From the formulation above, the ultimate goal is to reduce SEIR propagation on the post-intervention graph. Directly optimizing SEIR outcomes is expensive because it typically requires many running stochastic simulations. SPID therefore optimizes a structural surrogate: by breaking short source-to-target routes, diffusion is forced onto longer detours, reducing both speed and reach.

Given G, S_0, T_0, and budget B, SPID samples K representative source–target pairs $P = \{(s_i, t_i)\}_{i=1}^{K}$. Sampling can be performed in two modes:

Uniform: sample uniformly at random from $S_0 \times T_0$.
Risk-aware: sample according to weights $r(s)$ (e.g., source out-degree) and $v(t)$ (e.g., target community size). When such metadata are unavailable, uniform sampling is used. For each pair $(s, t) \in P$, SPID computes the shortest-path set $\mathrm{SP}(s, t)$ (or a single representative shortest path). Each candidate node v (excluding sources and targets) receives a score based on how frequently it appears on these paths, multiplied by a priority weight that favors earlier interdiction near the source (e.g., a decreasing function of distance from s).
To deal with rerouting after interdiction, we use SPID $++$, a greedy version of SPID with selective recalculation: at each step, it recomputes shortest paths only for the sampled pairs affected by the newly removed element, reducing unnecessary work.

Algorithm 1: SPID++

Input: $G = (V,E)$, S_0, T_0, budget B, sample size K.
Output: intervention node set U.

1. Sample K pairs $P = \{(s,t)\}$ from $S_0 \times T_0$ (uniform or risk-aware).
2. For each $(s,t) \in P$, compute $\mathrm{SP}(s,t)$ and initialize node scores.
3. $U \leftarrow \varnothing$.
4. For $b = 1,\ldots,B$:
   a. Select $u \in V \backslash U$ with the highest score; add u to U; remove u from G.
   b. Identify affected pairs $P_u = \{(s,t) \in P : u \text{ lies on the current path for } (s,t)\}$.
   c. Recompute $\mathrm{SP}(s,t)$ only for $(s,t) \in P_u$ on the updated graph.
   d. Update scores by subtracting the previous contributions and adding the new contributions induced by P_u.
5. Return U.

The computational bottleneck of SPID is repeated shortest-path computation and updates across many pairs. SPID addresses this by integrating an efficient SP-Kernel based on a blocked adjacency-matrix formulation for parallelization [24], further extended to multi-core execution with workload partitioning and synchronization for large graphs [25].

SP-Kernel supports batched initial shortest-path computation and selective updates for affected pairs, balancing interdiction effectiveness with practical runtime on large-scale networks. Initialization scales with K shortest-path queries. Each greedy iteration depends on the number of affected pairs (typically much smaller than K). In experiments, we use $K = 5000$, a distance-decay weight with $\alpha = 1$, and budgets $B = 10$–60 for sensitivity analysis over budget levels.

4 Experiments and Evaluation

This section describes the datasets, baselines, SEIR simulation settings, and evaluation protocol used to validate SPID under two different network regimes: hub-dominated and clustered small-world structures.

4.1 Datasets

The two networks considered here have different topological features for the analysis of the efficacy of the intervention strategies proposed. They are a real-world network and a synthetic network.

- **PolitiFact (FakeNewsNet):** The network used comes from the FakeNewsNet repository [7]. The network is characterized by scale-free properties, which means that the network follows a power-law degree distribution and has high-degree nodes. After

preprocessing, there are 2,500 nodes and 7,442 edges. Since scale-free networks rely heavily on their hubs in terms of diffusion of information, this network provides useful data to evaluate interventions that target hubs and analyze the behavior of SPID when the network is dominated by hubs.

- **Watts-Strogatz (Synthetic):** This is a synthetic network that is created using the Watts-Strogatz graph (N = 2,000). The graph represents a small-world network where the clustering coefficient is high, which makes the transmission process depend heavily on bridging edges.

4.2 Baselines and Parameter Settings

We evaluate SPID + + as the main implementation of the proposed SPID framework against four baselines: No Intervention, Random selection, Degree Centrality, and Betweenness Centrality. Unless otherwise noted, all methods are tested under the common configuration reported in Table 1.

Table 1. Experimental parameters.

Parameter group	Symbol	Value
SEIR	β	0.05
SEIR	δ (E $\rightarrow$ I)	0.25
SEIR	γ (I $\rightarrow$ R)	0.10
SEIR	T (steps)	30
SPID	K (sample pairs)	5000
SPID	α (distance weight)	1
Intervention	B (budget)	{10,20,30,40,50,60}
Sources/targets	$\lvert S_0 \rvert$	10 ($\approx$0.5% of node)
Sources/targets	T_0	$V \setminus S_0$
Monte Carlo	R	20

We chose $K = 5000$ to balance node-score stability and running time while keeping the SEIR parameters fixed across all methods to isolate the effect of the intervention strategy. Preliminary tests showed that smaller values of K led to unstable node rankings, while larger values only slightly changed the top-ranked nodes but significantly increased the shortest-path computation. Therefore, K functions primarily as a sampling-cost parameter rather than a diffusion parameter.

4.3 Evaluation Metrics

We evaluate mitigation performance using two measures that characterize outbreak severity and diffusion extent.

- **Peak Infected (Peak):** the maximum number of nodes in state I at any time during the simulation. This measure indicates the highest outbreak intensity and is closely related to short-term control pressure.
- **Total Infected:** the number of nodes that ever become infected during the simulation, corresponding to the cumulative affected population. This measure captures the final reach of the diffusion process.

Results for both metrics are presented across intervention budgets and network topologies.

4.4 Experimental Protocol

For each experimental setting – defined by the dataset, SEIR parameters, intervention budget B, and strategy – we run the stochastic SEIR simulation 20 times and report the average results. Repeating the simulation in this way helps reduce stochastic variation and keeps the comparison fair across strategies. In addition to comparing methods at a fixed budget, we examine budget sensitivity by varying B over a range of values. This analysis allows us to study how each strategy benefits from additional intervention resources and whether SPID scales more effectively than Degree Centrality or Betweenness Centrality across different network topologies.

5 Results and Discussion

This section reports the performance of SPID ++, the residual-updated implementation of SPID, in comparison with the baseline methods. We focus first on mitigation effectiveness in small-world networks represented by the Watts-Strogatz model, and then examine how performance changes as the intervention budget increases. The results are subsequently compared with those obtained on the real PolitiFact network in order to highlight how different network structures influence strategy behavior.

5.1 Performance on Small-World Networks (Watts–Strogatz)

In Watts-Strogatz networks, node degrees are relatively homogeneous, whereas clustering remains high. Under these conditions, cross-community diffusion is often driven by shortcut links and bridge nodes or edges. As a result, strategies based solely on degree may not identify the most influential transmission routes. By contrast, SPID is designed to target diffusion corridors approximated through shortest paths, which may provide an advantage in this type of topology.

Table 2 compares Peak and Total Infected at the same intervention budget (Table 3 and Fig. 3).

Random yields little improvement and, in some runs, slightly increases Total Infected, likely because the selected nodes vary substantially in importance. Degree Centrality reduces Total Infected by 4.3%, while SPID++ achieves a larger reduction of 9.5%, corresponding to about 2.2 times the improvement obtained by Degree in the small-world setting. Betweenness Centrality also performs reasonably well, with a

Table 2. Comparison of mitigation performance on the Watts-Strogatz network.

Strategy	Peak	Total Infected	% Reduction (Total)
No Intervention	397.6	1741.0	0.0%
Random	395.6	1764.0	-1.3%
Degree	375.4	1666.8	4.3%
Betweenness	360.9	1608.5	7.6%
SPID++ (Ours)	**351.9**	**1575.2**	**9.5%**

Table 3. Indicative CPU running time of intervention selection at $B = 60$.

Strategy	PolitiFact	Watts–Strogatz
Degree Centrality	0.01 s	0.01 s
Betweenness Centrality	1.42 s	1.18 s
SPID++ without batching	0.76 s	0.84 s
SPID++ with batching	0.41 s	0.46 s

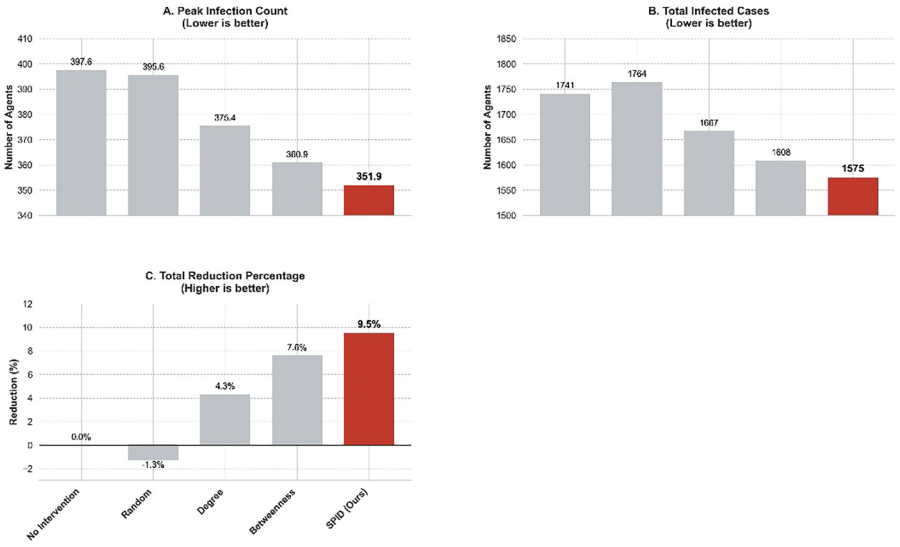


Fig. 3. Comparison of Mitigation Strategies on Watts-Strogatz Networks

reduction of 7.6%, which is consistent with its sensitivity to bridging structure. Even so, SPID++ performs best, likely because it targets diffusion corridors through source–target shortest-path sampling and updates the intervention set after each step.

5.2 Budget Sensitivity and Scaling Behavior

We further study budget sensitivity by varying the intervention budget B and measuring Total Infected for each method. Two consistent trends are observed. Degree Centrality saturates early: as B increases, the incremental reduction in Total Infected becomes smaller, suggesting that the most influential high-degree nodes are selected early and that further removals contribute relatively little to reducing inter-community transmission. SPID++, by contrast, continues to benefit from larger budgets. Its residual updating allows the method to adapt to rerouted diffusion paths after each intervention step, so additional budget can disrupt newly emerging transmission corridors.

At $B = 60$, SPID++ protects 444 more nodes than Degree Centrality, meaning that Total Infected is lower by 444. This result indicates that, in clustered networks, larger budgets are used more effectively when the intervention strategy targets diffusion corridors rather than relying only on node degree. We also report indicative CPU running times for the intervention selection step at the same budget level, where shortest-path computation is the main cost of SPID. Batching reduces the running time of SPID++ while keeping the selected intervention set unchanged.

These results show that SPID++ is more expensive than Degree Centrality, but remains faster than Betweenness Centrality; the batched version gives about a 1.8 $\times$ speedup on both tested graphs.

5.3 Comparison on the Real PolitiFact Network (Hub-Dominated)

On the PolitiFact network, which exhibits the hub-dominated characteristics typical of a scale-free topology, Degree Centrality is particularly effective at reducing Peak Infected, achieving a reduction of about 99%, whereas SPID++ achieves about 88%. This is consistent with scale-free intuition: disabling hubs can sharply suppress early-stage dissemination capacity.

Nevertheless, SPID remains highly effective and practically relevant in settings where direct hub intervention is infeasible (e.g., policy constraints, accounts that cannot be removed, or requirements for less intrusive actions). In such cases, SPID's corridor-interdiction mechanism serves as a complementary alternative: it reduces spread by cutting short routes from the source region to other clusters, rather than depending entirely on hub removal.

6 Conclusion and Limitations

We propose SPID, a corridor-based framework for mitigating harmful-information diffusion in social networks. SPID selects intervention nodes according to their coverage of shortest paths from source regions to target communities. In the experiments, we use SPID++, a residual-updated implementation of SPID that adapts to alternative routes after each intervention step.

Experiments show that SPID++ is particularly effective on high-clustering small-world networks. On Watts–Strogatz graphs, SPID++ reduces Total Infected more than Degree Centrality and maintains a clear advantage as the budget increases, with the

performance gap widening at higher budgets. On the real PolitiFact network, Degree Centrality is stronger in Peak reduction due to hub dominance, yet SPID++ still provides high mitigation and remains useful as a non-hub-dependent intervention strategy. For deployment, SPID++ can be paired with an efficient shortest-path kernel that supports batched shortest-path queries and improves feasibility on larger graphs.

This study assumes a static network per simulation episode, known initial sources, medium-sized graphs, and node-based interventions under a fixed budget. Future work will extend SPID to larger and dynamic graphs, partially observed source sets, and soft interventions, where transmission is attenuated rather than nodes or edges being fully removed.

Acknowledgment. This work was supported by the Ministry of Science and Higher Education (assignment No. FSEE-2025-0015).

References

1. Vosoughi, S., Roy, D., Aral, S.: The spread of true and false news online. Science **359**(6380), 1146–1151 (2018). https://doi.org/10.1126/science.aap9559
2. Del Vicario, M., Bessi, A., Zollo, F., et al.: The spreading of misinformation online. Proc. Natl. Acad. Sci. U.S.A. **113**(3), 554–559 (2016). https://doi.org/10.1073/pnas.1517441113
3. Shao, C., Ciampaglia, G.L., Varol, O., Yang, K.-C., Flammini, A., Menczer, F.: The spread of low-credibility content by social bots. Nat. Commun. **9**, 4787 (2018). https://doi.org/10.1038/s41467-018-06930-7
4. Törnberg, P.: Echo chambers and viral misinformation: modeling fake news as complex contagion. PLoS ONE **13**(9), e0203958 (2018). https://doi.org/10.1371/journal.pone.0203958
5. Shu, K., Sliva, A., Wang, S., Tang, J., Liu, H.: Fake news detection on social media: a data mining perspective. SIGKDD Explor. Newsl. **19**(1), 22–36 (2017). https://doi.org/10.1145/3137597.3137600
6. Zhou, X., Zafarani, R.: A survey of fake news: fundamental theories, detection methods, and opportunities. ACM Comput. Surv. **53**(5), Article 109 (2020). https://doi.org/10.1145/3395046
7. Shu, K., Mahudeswaran, D., Wang, S., Lee, D., Liu, H.: FakeNewsNet: a data repository with news content, social context, and spatiotemporal information for studying fake news on social media. Big Data **8**(3), 171–188 (2020). https://doi.org/10.1089/big.2020.0062
8. Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M., Procter, R.: Detection and resolution of rumours in social media: a survey. ACM Comput. Surv. **51**(2), Article 32 (2018). https://doi.org/10.1145/3161603
9. Pennycook, G., McPhetres, J., Zhang, Y., Lu, J.G., Rand, D.G.: Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. Psychol. Sci. **31**(7), 770–780 (2020). https://doi.org/10.1177/0956797620939054
10. Basol, M., Roozenbeek, J., van der Linden, S.: Towards psychological herd immunity: cross-cultural evidence for two prebunking interventions against COVID-19 misinformation. Big Data Soc. **8**(1) (2021). https://doi.org/10.1177/20539517211013868
11. Kim, J., Tabibian, B., Oh, A., Schölkopf, B., Gomez-Rodriguez, M.: Leveraging the crowd to detect and reduce the spread of fake news and misinformation. In: Proceedings of the 11th ACM International Conference on Web Search and Data Mining (WSDM 2018), pp. 324–332. ACM (2018). https://doi.org/10.1145/3159652.3159734

12. Tong, G., Wu, W., Du, D.-Z.: On misinformation containment in online social networks. In: Advances in Neural Information Processing Systems (NeurIPS 2018). arXiv:1809.06486 (2018)
13. Farajtabar, M., et al.: Fake news mitigation via point process based intervention. In: Proceedings of the 34th International Conference on Machine Learning (ICML 2017), Proceedings of Machine Learning Research, vol. 70, pp. 1097–1106. PMLR (2017)
14. Abouzeid, A., et al.: Socially fair mitigation of misinformation on social networks via constraint stochastic optimization. In: *Proc. AAAI Conf. on Artificial Intelligence* 36(11), 11801–11809 (2022). https://doi.org/10.1609/aaai.v36i11.21436
15. Simpson, M., Hashemi, F., Lakshmanan, L.V.S.: Misinformation mitigation under differential propagation rates and temporal penalties. Proc. VLDB Endow. **15**(10), 2216–2229 (2022). https://doi.org/10.14778/3547305.3547324
16. Xu, L., et al.: Selecting debunkers via self-imitation learning. In: Proceedings of the AAAI Conference on Artificial Intelligence **38**(20), 22447–22456 (2024). https://doi.org/10.1609/aaai.v38i20.30252
17. Smith, J.C., Song, Y.: A survey of network interdiction models and algorithms. Eur. J. Oper. Res. **283**(3), 797–811 (2020). https://doi.org/10.1016/j.ejor.2019.10.030
18. Holzmann, T., Smith, J.C.: Shortest-path interdiction problem with arc improvement recourse: a multiobjective approach. Nav. Res. Logist. **66**(3), 230–252 (2019). https://doi.org/10.1002/nav.21839
19. Punla-Green, S.Z., et al.: Shortest path network interdiction with asymmetric uncertainty. Networks **85**(5), 886–905 (2024). https://doi.org/10.1002/net.22208
20. Azizi, E., Seifi, A.: Shortest path network interdiction with incomplete information: a robust optimization approach. Ann. Oper. Res. **335**(2), 727–759 (2024). https://doi.org/10.1007/s10479-023-05350-1
21. Zhang, L., Chen, Z., Lu, C.-T., Zhao, L.: Network interdiction goes neural. arXiv:2405.16409 (2024)
22. Moug, K., Shen, S.: Stochastic bilevel interdiction for fake news control in online social networks. Comput. Oper. Res. **173**, 106872 (2025). https://doi.org/10.1016/j.cor.2024.106872
23. Sridhar, U., Blanco, M., Mayuranath, R., Spampinato, D.G., Low, T.M., McMillan, S.: Delta-stepping SSSP: from vertices and edges to GraphBLAS implementations. In: Proceedings of the 2019 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), pp. 241–250. IEEE (2019). https://doi.org/10.1109/IPDPSW.2019.00047
24. Shichkina, Y., Nguyen, XH., Ha, M., Tran, D.M.: Shortest path search method on a graph with cycles. In: Gervasi, O., Murgante, B., Garau, C., Taniar, D., C. Rocha, A.M.A., Faginas Lago, M.N. (eds.) Computational Science and Its Applications – ICCSA 2024 Workshops. ICCSA 2024. Lecture Notes in Computer Science, vol. 14825, pp. 344–356. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-65343-8_26
25. Ha, M., Shichkina, Y., Nguyen, X.H.: Extending the shortest path algorithm for large graphs with cycles and parallel computing capabilities. In: Gervasi, O., et al. Computational Science and Its Applications – ICCSA 2025 Workshops. ICCSA 2025. Lecture Notes in Computer Science, vol. 15894, pp. 191–204. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97648-3_13

New Digital Birth Certificate Based on Biometric and Documentary Data of the Child

A. D. Shchegolev(✉), G. A. Kukharev, and N. L. Shchegoleva

Saint-Petersburg State University, Saint-Petersburg, Russia
scetch2003@gmail.com

Abstract. This paper proposes a new format for a digital Birth Certificate. It contains a color photograph of the child's face and facial biometric characteristics, documentary information (child's full name, year of birth, etc.), information about the educational institution, and the full names of the mother and father. The developed software enables the creation of a digital Birth Certificate in the form of a BIO QR code that stores all the necessary documentary information about the child, which can be read by standard barcode scanners and software. The BIO QR code is embedded in the LSB layers of the color image of the child's face, allowing for the concealment of information, as the resulting image with embedded data is indistinguishable from the original. The proposed digital Birth Certificate is a GIF file consisting of a set of multimedia files (graphic, biometric, and documentary) that compactly store information about the child. The Digital Birth Certificate is a document that can be updated at various stages of a child's development and allows for the adjustment of documentary information and the characterization of their facial biometrics.

Keywords: digital birth certificate · QR code · BIO QR code · COLOR QR code · LSB layer · GIF file · biometric characteristic · biometric technologies · human phenotype

1 Introduction

In 2024, State Duma Deputy Sultan Khamzaev of the Russian Federation encountered a proposal to amend the form and content of the birth certificate. In her address to the Deputy Prime Minister, Deputy Tatyana Golikova proposed modernizing the document by adding photographs of the child at various stages of development (ages three, five, and seven). Furthermore, she proposed creating a new digital version of the certificate, which would be accessible through the government services portal. The deputy argued the need for such changes, as the document's format is inconsistent with modern digitalization and security requirements.

With the rapid development of digital technologies and the transition to electronic document management, the birth certificate form is being introduced. The lack of a digital version of the certificate and its administrator on digital platforms such as Gosuslugi makes it ineffective for identifying and recording minors. Implementing this new form

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 415–432, 2027.
https://doi.org/10.1007/978-3-032-30530-5_26

will allow the creation of a certificate that will simplify citizen interactions with the government and enhance the protection of personal data. One of the key issues with the current form of birth certificate is its inability to prevent children from being unexpectedly transported across borders. The lack of photographs of the child in the document allow criminals to use forged documents to transport children, posing serious risks to their safety. This is especially relevant given the increasing number of international marriages and divorces, when one parent may have a child without the consent of the other parent.

Based on the above, it can be concluded that developing a new type of birth certificate is essential. This would include both verbal information (a description of the child and the parents) and graphic information (photographs of the child at various stages of development). Furthermore, the document should be available in digital form, which would simplify the identification and accounting process and improve child safety. Implementing such changes would address issues related to the illegal removal of children, the search for missing children, and their identification, ultimately enhancing the protection of minors' rights. The introduction of photographs and the adoption of a document format allow volunteer organizations to more quickly and accurately locate missing children, speeding their successful recovery. The new document format must be developed with modern technology and security requirements in mind to ensure the protection of the interests of children and their parents.

1.1 The Origin and Development of the Identity Card

The development of the passport system as a mechanism for identification of a person began in Europe in the 14th century, when international travel for trade and other purposes required the introduction of identity documents. The first travel documents contained information about the holder: first name, last name, hair and eye color, and nose shape. These documents can be considered as the prototypes of modern international passports. In the 15th century, they began to be issued not only to residents visiting another country but also to foreigners who needed to travel within the country.

In Russia, the history of the passport began in the 17th century with the introduction of so-called "dangerous letters" for persons traveling abroad. With the introduction of the passport system in 1724, peasants began to be issued "pass letters" necessary for traveling. Forgery of such documents was severely punished, including hard labor. The 19th century saw significant changes. In 1833, the "Code on Passports and Runaways" was approved, regulating the issuance and registration of passports. The Statute on Passport Documents, developed in 1894, provided for three types of passports: temporary passport books, poster passports, and permanent passport books.

After the October Revolution of 1917, the Soviet government introduced work books as identification documents. In 1930s, the passport system has modernized, and new passports began to contain a photograph of the holder.

After the war, the passport system underwent further standardization. Decrees of 1953 and 1974 introduced a uniform procedure for issuing passports, mandatory for all USSR citizens over 16 years of age. Passports recorded a person's last name, first name, patronymic, date and place of birth, marital status, and other information. Passports in Russia have evolved from simple travel documents to a comprehensive system of

personal identification, fulfilling the functions of monitoring, recording, and protecting citizens' rights.

However, it is worth noting that passports were and are currently issued to adults, while birth certificates are used for children. Before the Revolution, children born to commoners were registered in church parish registers; later, birth registers were introduced, which recorded the birth and issued a corresponding certificate. The document was called an "Extract from the Birth Register," popularly known as a "metric." In 1935, a decree of the Council of People's Commissars of the USSR approved a standard form for birth certificates and made their issuance mandatory. Currently, in the Russian Federation, the content of the Birth Certificate is regulated by Article 23 of Federal Law No. 143 of November 15, 1997.

1.2 Problems of Children's Travel Abroad

A significant number of children go missing in Russia every year. According to statistics, approximately 40,000 children disappear annually, and only a fraction of them are ever found [1–3]. The current birth certificate format creates problems when crossing the border, when proof of legal guardianship (parents) is required, when registering at educational institutions, and in other situations where confirmation of a minor's identity is required. This is especially true for children who are unable to provide information about themselves or for those separated from their families due to emergency situations.

Currently, the problem of illegal transporting children abroad is becoming increasingly pressing. This practice not only violates the rights of children and their legal guardians but also creates serious threats to the safety and well-being of minors.

According to international and national law, removing a child abroad without the consent of one of the parents or legal guardians violates the rights of the child and their legal guardians. In particular, the Hague Convention on the Civil Aspects of International Child Abduction of October 25, 1980, establishes mechanisms for the return of children illegally removed abroad. However, the use of false documents significantly complicates the process of identifying and preventing such actions.

Forged documents used to remove children may include false passports, visas, travel consents from the other parent or guardian, and false birth certificates. Such documents are often produced with a high degree of authenticity, making them difficult to detect at border control. As a result, children may be taken to another country, where their whereabouts become difficult to determine.

1.3 Social and Psychological Consequences

The illegal removal of children abroad using false documents has serious social and psychological consequences for both the children themselves and their families. For a child, such a transfer can be a traumatic experience, especially if they find themselves in a foreign country without the support of loved ones. Children may face language barriers, culture shock, and a lack of access to education and healthcare.

For parents and legal guardians, the loss of a child is a difficult emotional ordeal. Searching for and returning a child can take years, leading to lengthy legal proceedings

and significant financial costs. Furthermore, such situations are often accompanied by conflicts between parents, which negatively impacts the psychological climate in the family.

1.4 Measures to Prevent the Removal of Children with Forged Documents

To prevent children from being taken abroad using forged documents, comprehensive measures are needed, including both tightening border controls and improving legislation. One key area is the implementation of modern technologies, such as biometric identification systems, which allow for more effective detection of forged documents when taking children abroad.

Biometric identification is the process of identifying an individual based on a person's unique physiological or behavioral characteristics [4–10]. Unlike traditional authentication methods, biometric data cannot be forgotten, lost, or transferred to another person, making it more reliable and convenient for identifying children when taking them abroad. Facial image identification, the most common biometric authentication methods, appears to be the most promising. This method is based on the analysis of unique facial features, such as the distance between the eyes, the shape of the nose, cheekbones, and other parameters. However, if a person's appearance changes, this can lead to errors, meaning an incorrect decision will be made that the person is not the one depicted on the document.

This is precisely the main problem with identifying children—their appearance changes dramatically over a short period of time (sometimes less than a year). The shape of the nose, the contour of the face, the relationships between facial features (eyes, nose, mouth), and so on change. It is believed that significant changes in a person (not related to injury or surgery) end around age 25. However, these changes are most noticeable in children under 14–15 years of age. Therefore, issuing children a passport that is valid for 5–10 years and contains only one image at the time of issue creates problems during identification at customs control. A solution to this problem could be a new version of the electronic Birth Certificate, which will contain not only documentary information about the child, but also a current image of the child, as well as biometric characteristics, allowing for significantly more accurate identification, thereby preventing the illegal removal of children outside the Russian Federation.

1.5 Ways to Solve the Problem

It should be noted that border crossings typically require document checks for a large number of people. Adding a large amount of data to enable more precise identification can increase the time it takes inspectors to complete the check. Therefore, to optimize and accelerate this process, it is necessary to implement a method that will make this check as quick as possible. An obvious solution is to use barcoding methods for storing and transmitting data, which are widely used in various fields [12].

The idea of using barcodes to identify people during electronic payments was first presented in 1999 in [13]. In 2009, works [15, 16] appeared that proposed describing all information about a person's face using a set of black and white horizontal lines (brow line, eye line, and lip line), thus forming a binary code, which the authors called a

"biological barcode." Later, researchers [17] noted that transforming a biological barcode into a standard one would allow its use for identifying people. However, the authors were unable to propose approaches that would generate such a barcode.

Methods for constructing and applying a biological barcode for comparing facial images were first presented in the monograph [5]. The monograph proposed an algorithm closest to the concept of a linear "biological barcode," based on the use of a gradient in a sliding window. The problem with this approach is the impossibility of obtaining a stable code for facial images of the same person if they have different characteristics (e.g., different facial sizes in the images, head tilts, or emotions).

The next step in the development of methods for generating a barcode based on human biometric characteristics (facial expression) was the publication of works [6, 7]. The method proposed in [6] is based on constructing a luminance histogram for a grayscale image generated from the original color image, then averaging the histogram over a limited number of luminance intervals, quantizing the resulting values in the decimal range from 0 to 9, and converting the resulting data into a linear EAN-8 barcode using a standard algorithm. This approach also has some limitations, but is more stable than the one discussed previously.

The barcode generation method proposed in [7] and formalized in [8] involves calculating the distance between two identical mirror-image windows on a facial image, which move downwards in increments from the hair/forehead boundary to the lower boundary of the nose. The luminance gradients of this image are then calculated. The resulting gradients are then normalized, discretized, and the resulting values are averaged. To generate a standard image barcode, the resulting values are quantized to a decimal range from 0 to 9. The study demonstrated that the resulting barcode is stable under changes in the local dimensions of faces, their rotation, angle changes, and mirror rotation around a vertical axis, as well as in the presence of facial expressions and age-related changes. However, this method does not require significant computational effort or specialized image processing software. Since protecting children's facial images requires combining them with biometric facial characteristics and presenting them together in a single "digital portrait," it is necessary to study methods for embedding documentary and other information into barcodes.

1.6 Overview of Solutions that Use QR Codes in Personal Documents

The use of barcodes has become an effective tool during the COVID-19 pandemic, allowing you to track the vaccination status of citizens. QR TIGER [18] proposed to use QR codes for the Health Passport, storing information about human health in it. The passport contains medical history, information about allergies, the need to take medications, etc. Scanning a QR code with a smartphone allows you to access all medical information about a person. This makes it easier to transfer, share, and store important health information.

However, this QR code does not allow you to identify a person and is essentially only a way to access information about him. The same can be said for the Multimedia QR Code proposed [12] – the QR code stores only the URL that redirects the user to a landing page or a location in the cloud where the media file or PDF, PNG, JPEG, MP4 file, or Excel spreadsheets and Word files are stored.

The first work devoted to embedding biometric information in barcodes was [9]. This paper provides a list, characteristics and volume of facial biometric information used for bar coding. It is shown that the criteria for choosing a barcode for storing documentary, graphic and biometric information about a person are: sufficiency of color images of faces and anthropometric points for recording and storing information; the ability to write: numbers, special characters, numerical matrices, alphabetic information in Cyrillic and Latin; maximum geometric dimension; ease of writing and reading information and the popularity of using a barcode.

Only two-dimensional barcodes most fully meet these criteria, namely GS1 QR code, which has a capacity of 7089 digits or 4296 alphanumeric characters. At the same time, it can be used to build not only a color QR code [5–7, 20], but also a color barcode, which has the structure of halftone and color images [21, 22].

At the same time, several layers of QR codes, combined into a common layered structure, form a colored QR code [20], which has a data volume three times larger than a standard one. Such a color QR code is essentially an image and can be recorded in RGB scale, and the use of the.png (portable network graphics) format allows you to resize the image without losing information when reading from the color QR code.

The idea of creating and reading a color QR code is presented in Fig. 1. However, the use of such QR codes can reduce the likelihood of detecting a read error, as well as waste of color image resources, since all three images store the same message.

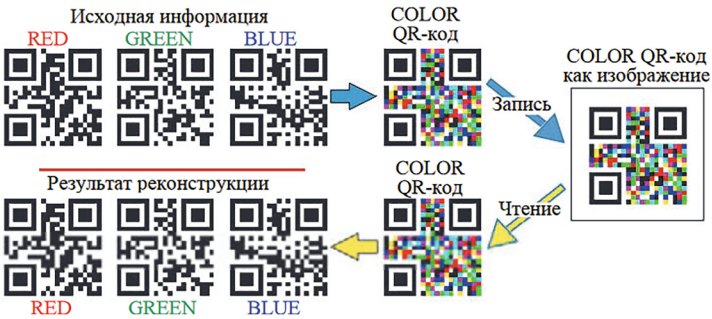


Fig. 1. Idea of creating and reading a color QR code [22]

Thus, you can modify this approach by putting different information in different layers of the image, and use a color image of a face as a drawing. This served as the basis for the approach to the formation of the BIO QR code proposed in [9]. Here, in order to obtain a BIO QR code, biometric information is prepared, which must be placed in the code. This information is recorded in the form of a QR code and embedded in the color image by replacing its G component (see Fig. 2). However, this substitution makes it impossible to obtain a color image of the face, since the image will be incorrect when merging the layers.

In [11], a method is proposed for embedding three different QR codes with biometric information about a person into pre-zeroed virtual LSB layers of the Red, Green and Blue components of a color face image. The size of each QR code must be equal to

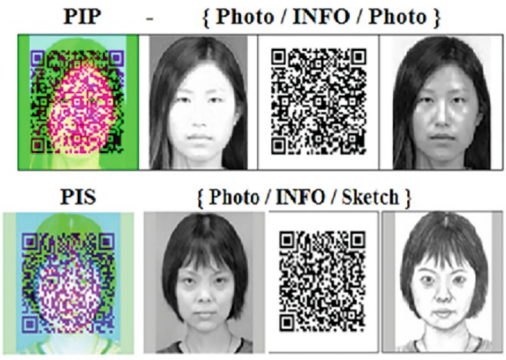


Fig. 2. PIP and PIS options for QR code [9]

the size of the original face image. In the paper [23], an example of using an animated biometric QR code is given to implement the procedure of control and identification of people by placing it on passes. This QR code, despite the fact that it contains a minimum amount of personal information, can provide identification of a person.

A number of works are devoted to hiding information stored in QR codes in order to restrict access to it. For example, [24] proposes a hybrid scheme based on a QR code and a discrete wavelet transform (DWT), which makes it possible to protect information. In [25], it is proposed to encrypt information before placing it in a QR code. In [26], a QR code is used as a container, and the information contained in it can be read only if there is a secret key generated by the authors' algorithm. However, despite the presence of a large number of bar coding technologies for identifying people, none of them fully satisfies the implementation of a modern birth certificate on its basis.

1.7 Requirements for the Content of the Birth Certificate

Facial image recognition is becoming one of the key elements of the digital identity of an "electronic citizen", therefore it is necessary to present all facial data in a form that ensures their protection from unauthorized use, modification or substitution and transmission without loss of information.

The existing Birth Certificate has lost its relevance, so it is necessary to develop an option for submitting a new digital Birth Certificate for a child, which should include:

a color photo of the child's face and the characteristics of his facial biometrics - phenotype (face shape, hair and eye color, skin color, special features); documentary information - surname, name, patronymic of the child, date and place of birth, gender; information about the educational institution (if attending); Name, patronymic, surname and color photos of his parents.

All documentary information contained in the digital Birth Certificate must be read by standard scanners and software. Such certificate should be a supplementable document at different stages of a child's development (e.g., at the age of three, seven, and ten) with the ability to correct documentary information and facial biometrics. The most promising is the use of colored biometric QR codes and the storage of information about the child inside his image in LSB layers.

A child's digital birth certificate must be represented by a set of multimedia files (graphic, biometric and documentary) that compactly store the necessary (or all of the above) information with free or secure access to all data.

Graphic multimedia files must represent: images of faces, binary and color QR codes, BIO QR codes, and multi-component QR codes (with mixed graphics) presented in.gif format. The original image of the child's face should be used as an empty container, in which colored QR codes with messages representing biometric or documentary data are embedded in them. Biometric multimedia files should consist of binary and color QR codes containing phenotypic data about the child's face (if necessary, physical, anthropometric characteristics can be added). Documentary multimedia files are binary and color QR codes containing real data about the child (gender, full name, date and place of birth, special features (face shape, hair and eye color, skin color), as well as documentary information about the child's parents. All media files should be able to store and transmit the information attached to them. At the same time, reading the contents of all multimedia files and a hard copy (paper document) should be performed in standard ways.

2 Description of Technologies for the Formation of a New Digital Birth Certificate for a Child

2.1 Biometric BIO QR Code

A digital image is a rectangular matrix consisting of M rows and N columns, each element of which contains the color of a pixel [27]. For color images, several components are used, and the color of the pixel is represented by the corresponding number of numerical values. Blue. Then the original size of the color image of the child's face is determined by three parameters: M × N × 3, where M and N are the number of rows and columns of the image, respectively.

According to the requirements of GOST [27], the size of the face image should be 320,240 × pixels, while the maximum barcode size is 177,177 × pixels GOST [28]. The barcode is added by 2 times, and the face image is supplemented with stripes of the appropriate size: a strip of 17 × 240 pixels is added at the top and bottom, and a strip of 354 57 pixels is added to the left and right.

2.2 Formation of an Information QR Code

To generate an information QR code containing documentary data about the child, a text message is prepared in the following format:

surname, first name, patronymic of the child//year of birth//full name of the mother//full name of the father.

If necessary, the content of the message may be changed at the request of the relevant state authorities.

A binary QR INFO code is generated for this message. For images of the child's face and the image of the QR code, the.png format is used, which allows you to realize the presentation of the QR code without compression noise, ensure the complete preservation

of information during memory operations and the absence of distortion of the information recorded in the QR code. Example of an informational QR code QR INFO generated by a message.

Shchegolev Aleksandr Dmitrievich//2…//Shchegoleva N… L…//Shchegolev D… Etc…. Shown in Fig. 3.

Fig. 3. Example of an informational QR code

2.3 Generation of a Biometric BIO QR Code

To generate a biometric BIO QR code in accordance with [11], it is necessary:

decompose the prepared image of the child's face (Sect. 2.1.1) into components R, G and B, so that: R = I(:,:,1); G = I(:,:,2); B = I(:,:,3), where I is a color image of the child's face with a size of 354 × 354 × 3; generate a color QR code by replacing the G layer with a 2× enlarged image of the generated QR INFO information code so that:

$$QR\ (:,:,1) = R;\ QR\ (:,:,2) = QR_{INFO};\ QR\ (:,:,3) = B.$$

An example of the obtained biometric BIO QR code obtained on the basis of the initial image of the child's face and the generated information QR code is shown in Fig. 4.

Fig. 4. Example of a biometric BIO QR code and its component

2.4 Embedding Biometric Information in LSB Layers

The use of a biometric BIO QR code in a digital birth certificate allows you to quickly read information about the child, but at the same time makes it available for use by attackers and can be replaced by them. Therefore, the amount of information stored in such a QR code is limited and contains only the full name of the child, his year of birth and the full name of his parents.

To increase the representativeness of information about the child, its security and universality, as well as the reliability of storage, it is necessary to embed biometric information in the color PNG image of the child's face. The requirement to provide graphic information only in PNG format is connected with the fact that it is needed to ensure that information is preserved in the LSB layer.

Steganography can be used. In this paper, it is proposed to use binary layers of color images to embed QR codes with information about the child. At the same time, images containing such QR codes remain without visual changes, retaining the possibility of subsequent data extraction. This can be done most effectively by placing the necessary information in LSB layers (LSB – Least Significant Bits), which will ensure that it is protected from unauthorized access.

2.5 Preparation of QR Codes with Biometric Information and Information About the Educational Institution

To embed in a color image, you need two QR codes with biometric information and information about the educational institution.

To generate ***a QR code with biometric information***, a text message is prepared with a description of the child's phenotype, which has the following format (instead of an ellipsis, a description corresponding to the child's face is inserted):

Surname Name Patronymic of the child//Year of birth//Age//Eye color: ...//Hair color: ...// Skin color: ...//Oval of the face: ...//Nose: ...//Ears: ...

If necessary, the content of the message can be changed.

This message generates a binary QR code$_{\text{QR PHENO.}}$ In this case, to compose a color QR code, you need to use the byte (halftone) format of QR codes in.png format. Such a recording takes up less memory space than recording in.jpg format, as shown in [29]. This method allows you to ensure the presentation of byte QR codes without compression noise, the preservation of information during memory exchanges and the absence of distortion of information recorded in the LSB layers of the QR code.

To generate a QR code with information about the educational institution in which the child is studying, a text message is prepared in the following format: surname first name patronymic of the child//year of birth//age//type of educational institution//educator (or class teacher)//phone number of the educator (or class teacher)//phone number of parents; Based on the prepared message, a QR code PASS is generated in a similar way to the QR code$_{\text{PHENO.}}$

2.6 Algorithm for Accessing Binary Layers of Color Images

A halftone image is a matrix A of size $M \times N$ with positive integers that determine the brightness of the pixel. The algorithm for accessing binary layers of a color image is based on the procedure for calculating the remainder by dividing all the values of the matrix A of size $M \times N$ by base 2 [11 $\times$], which will have values of 1 only in those pixels that correspond to odd values in matrix A, and the rest will have values equal to 0.

A color image is represented by a matrix C of size M × N × 3, consisting of three components – R, G, B, each of which has the size M × N. As a result of the calculation D = mod(C, 2), we get an image matrix D of size $M \times N \times 3$ containing 1 in those pixels that correspond to odd values in each layer of the matrix C and zeros in the remaining pixels. Each component of a color image can be decomposed into 8 "virtual bit slices" containing only zeros and ones, each of which corresponds to its own multiplier: 128, 64, 32, 16, 8, 4, 2, and 1. A multiplier of 1 corresponds to the last, Least Significant Bits (LSB). The algorithm for accessing binary layers of a color image is described in [29]. An example of decomposing a color image of a child's face into binary layers is shown in Fig. 5.

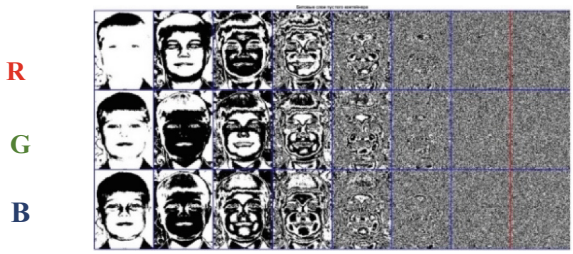


Fig. 5. Decomposing a color image into binary layers

Embedding QR codes in an image is based on "zeroing the LSB layers" of components R, G and B of a color image and placing the corresponding QR codes in the zeroed LSB layers [29]. Figure 6 shows LSB layers of a color image of a child's face with QR codes embedded in them. The remaining layers 2–8 do not change.

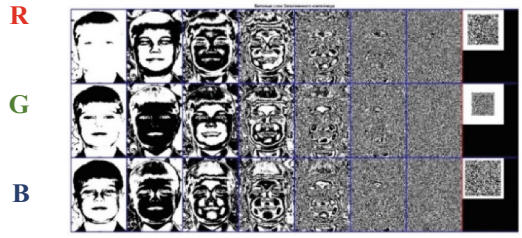


Fig. 6. LSB layers of a color image of the child's face with QR codes attached to them

The scheme of embedding QR codes in a color image of a face in accordance with the one proposed in [11] is shown in Fig. 7.

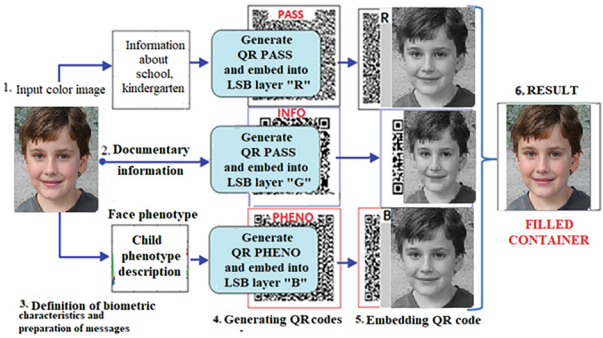


Fig. 7. Scheme for embedding QR codes in a color face image

The methods described above allow you to use the biometric characteristics of the child's face and supplement them with documentary information about him, as well as store this information in standard QR codes embedded in the LSB layers of the R, G and B components of the input color image of the face. At the same time, the resulting image with embedded QR codes containing information is a color container image that compactly stores all facial biometric and documentary information about the child.

This approach increases the security, universality and safety of information, which will ensure the restriction of open access to it, as well as make it possible to exchange information according to standards within the framework of international cooperation without losing any of its part and/or changing it. It is worth noting that QR codes embedded in a color image are not visually detectable, and information about the child can only be read and deciphered by specialists who have an idea of the rules for generating messages for generating QR codes. If it is needed to protect the information, the approach, suggested in [24], can be used, where the information is being encrypted first, then the QR codes are formed.

2.7 Submission of Birth Certificate Information in GIF Format

The GIF (Graphics Interchange Format) format was developed in 1987 at CompuServe. The structure of a GIF file depends on the version (GIF87a or GIF89a). In either version, the GIF file has a 13-byte header at the beginning that contains a signature that lists the GIF file type, GIF version number, and other information. To store images, the GIF format uses an indexed palette that has a palette of 256 colors. Each image in the file is accompanied by its own color table [30].

GIF files have the following advantages: provide lossless compression; GIF format is supported by all modern web browsers and operating systems (Windows, macOS, iOS, Android, etc.); viewing GIF files is possible on almost any modern device (PC, smartphone, tablet); GIF files are not modified when transferred to the server, which ensures the immutability of the information contained in them; Widespread use of this type of file.

The restrictions are: if an excessive number of images are used, it can reach a large size; GIFs need to be optimized, but a number of effective optimization algorithms have been proposed; for editing, you need to use the functions of correct reading and

formation of a new GIF file; a color palette of 256 colors (this is enough to represent the information of a digital birth certificate).

It is worth noting that GIF files are a set of images that appear for viewing in turn, which allows you to see all the information [10]. It is this feature that determines the choice of this format for the child's digital birth certificate, namely the ability to store in it not only documentary data in the form of a BIO QR code, but also photos of the child himself at the age at which he crosses the border, unlike the photo that is in the passport. Note that a biometric passport is issued for 10 years, and during this time the child's appearance and life circumstances can change significantly: for example, the child's hair or eye color changes, or due to the death of parents, he or she has a guardian or adoptive parents, the child can move to another educational institution, new special signs will appear, etc.

To streamline the information entered into the Digital Certificate, it is proposed to use the following template (if necessary, it can be changed or supplemented):

A photo of the child (for example, current or taken when crossing the state border for the first time).

Photo of the mother (guardian, legal representative, adoptive mother)
Photo of the father (guardian, adoptive father)
For age 1 (e.g. 1 year):
a photo of a child aged 1 year;
BIO QR code for age 1 year.
For age 2 (e.g. 3 years):
a photo of a 3-year-old child;
BIO QR code for age 3.
. . .
For age L (e.g. 17 years):
Photo of a child aged 17;
BIO QR code for age 17.

An example of the described content of the Birth Certificate file in GIF format is shown in Fig. 8.

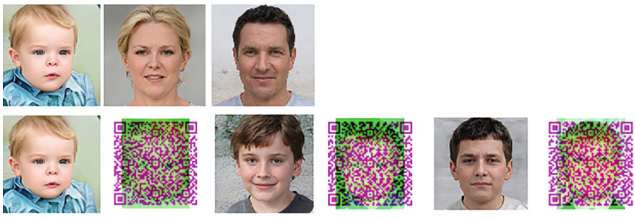


Fig. 8. Example of possible contents of a Birth Certificate file in GIF format

Presentation of the Birth Certificate as a GIF file allows you to store and view the information contained (biometric and documentary data about the child and his/her

parents (guardians, legal representatives) in it on any medium (smartphone, tablet, PC), using almost any operating system (Windows, macOS, iOS, Android, etc.).

3 Results

The developed application accepts as input a set of graphic, documentary and biometric information about the child and his parents.

XExample:

Documentary data

Surname: Shchegolev.

Full name: Alexander.

Fathername: Dmitrievich.

Year of birth: 2003.

Mother: Shchegoleva N… L…

Father: Shchegolev D… And…

Educational institution: SBEI Secondary School "Academic Gymnasium No56"// class teacher: T… Oh… N…// Phone: 322223322// Parents' phone: + 7.

Biometric data

Eye Color: Green// Hair Color: Blonde// Skin Color: White// Facial Oval: Round// Nose: Straight// Ears: Musical.

Graphic data

Photos of a child aged 2, 5, 9 years and photos of parents are given in Fig. 9.

Fig. 9. Sample Input Graphics

The process of generating a BIO QR code based on the entered documentary data and the image of the child is shown in Fig. 10: for age 2 years 5 years 9 years. The content of QR codes for different ages is distinguished by the explicit indication of the child's age in the photo.

An example of embedding information in LSB, layers of the child's face, QR codes with documentary and biometric information recorded in them are shown in Fig. 11. The information objects that are formed for the digital birth certificate of a child are shown in Fig. 12.

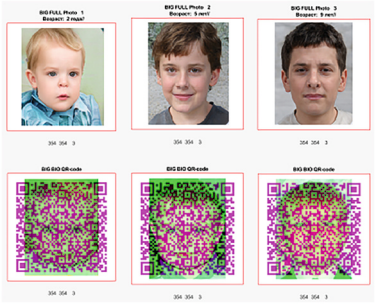


Fig. 10. Biometric QR code and BIO QR code components for child's face image (age 2 years, 5 years, 9 years)

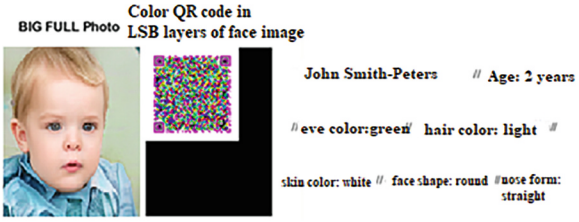


Fig. 11. Color QR code embedded in LSB layers of the child's face image, containing documentary and biometric information (age 2 years)

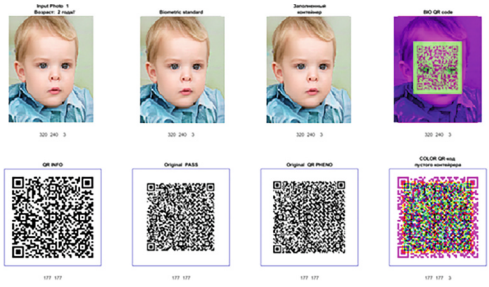


Fig. 12. Information objects that are formed for the digital Birth Certificate of a child (age 2 years)

A GIF file of a digital birth certificate for a child aged 9 years consists of the images shown in Fig. 13.

Fig. 13. Content of the GIF file of the child's birth certificate (age 9 years)

It should be noted that the content of GIF files for different ages of the child is different - for each next trip of the child outside the Russian Federation, a photo of the child and a BIO QR code corresponding to his current age are added. At the same time, an image of the child's face is generated with QR codes embedded in the LSB layers of the child's face image containing up-to-date documentary and biometric information, which is stored in the databases of state bodies (for example, in the Federal Migration Service (FMS of Russia).

In this case, the current age of the child is entered, and it is also possible to enter all the necessary data that have changed: appearance and special features, educational institution, change of surname, appearance of guardians or legal representatives of the child, etc.

4 Conclusion

1. The analysis showed that in order to protect children from illegal export abroad, a document is needed that includes not only documentary information, but also biometric information of the child, which will allow him to be accurately identified, which will increase the level of safety of children.
2. In this work, a new format of the child's digital birth certificate is proposed, which presents a color photo of the child's face and the characteristics of his facial biometrics, documentary information (full name of the child, year of birth, etc.), information about the educational institution, as well as the full names of the mother and father.
3. Software has been developed that allows you to generate a digital Certificate, which is a BIO QR code that stores all the necessary documentary information about the child. All documentary information contained in a digital birth certificate can be read by standard barcode scanners and software.
4. The developed software allows you to embed documentary and biometric information in the form of QR codes into LSB layers of a color image of the child's face for storage in the databases of government agencies, which allows you to hide the information, since the resulting image with embedded data does not differ from the original.
5. Software has been developed that allows you to generate and view a digital Birth Certificate of a child in a GIF file, consisting of a set of multimedia files (graphic, biometric and documentary), compactly storing information about the child.
6. A digital birth certificate is a supplementable document at different stages of a child's development and allows you to correct documentary information and characterize his/her facial biometrics.

References

1. Novoselova, E.: Every year, more than 40 thousand children disappear in Russia. https://rg.ru/2022/05/25/kazhdyj-god-v-rossii-propadaet-bolee-40-tysiach-detej.html
2. Pato, E., Rostovtseva, N.V.: Abduction of children: European and Russian prospects. J. High. Sch. Econ. **3** (2014). https://cyberleninka.ru/article/n/pohischenie-detey-evropeyskie-i-rossiyskie-perspektivy. Accessed 30 Mar 2025
3. Aslaeva, A.A., Gaevskaya, N.B.: Problems related to the illegal export of children from one country to another **3**(34) (2017). https://cyberleninka.ru/article/n/problemy-svyazannye-s-nezakonnym-vyvozom-detey-iz-odnoy-strany-v-druguyu. Accessed 30 Mar 2025
4. Frolova, E.Yu., Koshlykova, Yu.: Human identification by biometric data: a review of modern technologies. N. Caucasian Legal Bull. North Caucasus Legal Vestnik. **3**, 167–174 (2022)
5. Kukharev, G.A., Kamenskaya, E.I., Matveev, Yu.N., Shchegoleva, N.L.: Metody obrabotki i priznaniya imobrazheniya litsa v zadach biometrii [Methods of processing and recognition of facial images in biometric problems]. St. Petersburg, Polytechnic Publishing (2013). 388 p.
6. Kukharev, G.A., Matveev, Yu.N., Shchegoleva, N.L.: Express method of barcode generation based on images of persons. **2**(90), 99–106 (2014)
7. Kukharev, G.A., Matveev, Y.N., Shchegoleva, N.L.: Formation of barcode based on facial images based on brightness gradients. Sci. Tech. Bull. Inf. Technol. Mech. Opt. **3**(91), 88–95 (2014)
8. Kukharev, G.A., Shchegoleva, N.L.: Method of forming a barcode from images of faces and a device for its implementation – 2015 2015/2/27 Patent Office RU Patent number 2542886 Application number 2013159027/08, 30 December 2013
9. Kukharev, G.A., Kazieva, N., Tsymbal, D.A.: Barcode technologies for facial biometrics: current state and new solutions. Sci. Tech. J. Inf. Technol. ITMO Sci. Technol. Bull. **18**(1), 72–86 (2018)
10. Kukharev, G.A., Kazieva, N.: QR codes with animation for digital passes. Sci. Tech. Bull. Inf. Technol. Mech. Opt. **20**(3), 461–464
11. Kukharev, G.A., Kazieva, N., Shchegoleva, N.L.: Method of generating a color QR code based on images of faces and a device for its implementation– 2019 2019/7/2 Patent Office RU Patent number 2714741 Application number 2019124375/08, 30 December 2013
12. Top 40 innovative uses of QR codes. https://www.qrcode-tiger.com/ru/40-best-innovative-uses-of-qr-codes
13. Method for verifying human identity during electronic sale transactions – Patent US 5878155A (1999). [Electronic resource]. – Mode of access: http://www.google.com/patents/US5878155, free. Yaz. Eng
14. Barcode / Tattoos [Electronic resource]. – Mode of access: http://www.barcodeart.com/store/wearable/tattoos/, free. Yaz. Eng
15. Dakin, S.C., Watt, R.J.: Biological «bar codes» in human faces. J. Vis. **9**(4), 1–10 (2009)
16. Goffaux, V., Dakin, S.C.: Horizontal information drives the behavioral signatures of faces processing. Front. Psychol./Percept. Sci. **1**, 1–14 (2010). Article 143
17. Facial barcodes help us identify people. Mode of access: http://www.barcodesinc.com/news/?p=92, free. Yaz. Eng
18. Ways to use QR codes for health passport systems. https://www.qrcode-tiger.com/ru/qr-codes-for-health-passport
19. Multimedia QR code: easy way to share files. https://www.qrcode-tiger.com/ru/multimedia-qr-code
20. Duda, J.: Embedding grayscale halftone pictures in QR codes using correction trees. arXiv: 1211.1572v3[cs.IT], pp. 1–16, 2 December 2012

21. Garateguy, G., Arce, G.R., Lau, D.L., Villarreal, O.P.: QR images: optimized image embedding in QR codes. IEEE Trans. Image Process. **23**(7), 2842–2853 (2014)
22. Nesson, C.: Encoding multi-layered data into QR codes for increased capacity and security. Research Experience for Undergraduates, Summer (2013). Final - Report4.pdf
23. Ramesh, M., Prabakaran, G., Bhavani, R.: QR-code image steganography. In Proceedings of the second International Conference on Emerging research in Computing, Information, Communication and Applications (2014)
24. Shchegoleva, N., Zalutskaya, N., Dambaeva, A., Kiyamov, J., Dik, A.: New technologies for storing and transferring personal data. In: Gervasi, O., Murgante, B., Misra, S., Rocha, A.M.A.C., Garau, C. (eds.) Computational Science and Its Applications – ICCSA 2022 Workshops. ICCSA 2022. Lecture Notes in Computer Science, vol. 13380, pp. 680–692. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-10542-5_47
25. Alajmi, M., Elashry, I., El-Sayed, H.S., Farag Allah, O.S.: Steganography of encrypted messages inside valid QR codes. IEEE Access **8**, 27861–27873 (2020). https://doi.org/10.1109/ACCESS.2020.2971984
26. Baldin, M.I.: Analysis of the influence of color spaces on the results of color image processing by equalization algorithms **32**(3), 444–451 (2019). https://doi.org/10.15827/0236-235X.127.444-451
27. GOST R ISO/IEC 19794-5 – 2013 Information Technology. Biometrics. Biometric data exchange formats. Part 5. Facial image data
28. GOST R ISO/IEC 18004—2015 Information technologies. Automatic identification and data collection technologies. QR Code barcode symbology specification
29. Kazieva, N.: Methods and algorithms of bar coding for facial biometrics. Dissertation for the degree of Candidate of Technical Sciences. St. Petersburg (2020)
30. Demin, A.Y.: Fundamentals of Computer Graphics: A Textbook / A.Yu. Tomsk Polytechnic University. Tomsk Polytechnic University Publishing, Tomsk, 191 p. (2011)
31. Gonzalez, R.S., Woods, R., Eddins, S.: Digital image processing in the MATLAB environment. Chepyzhov, V.V. (ed.) Translation. Technosphere, Moscow (2006). 616 p., ISBN: 5-94836-092-X, 0-13-008519-7

Identifying the Characteristics of Brain Regions in Individuals with Different Neurocognitive Characteristics Based on Functional MRI Data

Nadezhda Shchegoleva(✉) and Kakabayeva Nurjemal

Saint Petersburg State University, St. Petersburg, Russia
n.shchegoleva@spbu.ru, st098457@student.spbu.ru

Abstract. Autism spectrum disorders (ASD) do not have objective neurobiological markers, and diagnosis is based solely on behavioral scales. Functional resting MRI allows you to study disorders of brain connectivity, but existing approaches face problems of high dimensionality of data, overfitting and low interpretability. The article proposes a methodology for identifying stable, interpretable biomarkers of functional connectivity, reliably distinguishing individuals with ASD and neurotypical controls. To reduce dimensionality and ensure interpretability, two-stage feature selection (t-test with FDR correction, SelectKBest) and logistic regression are used. The model achieved an accuracy of 71.8% (AUC = 80.5%, Sensitivity = 74.4%, Specificity = 69.6%) and identified 22 stable associations. Mapping revealed a predominant localization in the DMN with pronounced hypoconnectivity (17↓ / 4↑), as well as hyperconnectivity of subcortical structures (6↑ / 3↓) and the visual network (2↑).

Keywords: autism spectrum disorder · functional connectivity · resting-state fMRI · machine learning · brain networks

1 Introduction

Autism spectrum disorders (ASD) are a heterogeneous group of neurodevelopmental conditions characterized by deficits in social interaction and communication, as well as restricted and repetitive patterns of behavior, interests, and activities [1]. Despite considerable progress in understanding the genetic and neurobiological mechanisms underlying ASD, clinical diagnosis still relies primarily on behavioral criteria, including standardized scales and expert assessments. The subjectivity of these methods, their dependence on the clinician's experience, and the inability to detect the disorder before the onset of behavioral symptoms underscore the urgent need for objective, quantitative biomarkers.

Resting-state functional magnetic resonance imaging (rs-fMRI) has emerged as a powerful tool for investigating the functional architecture of the brain. Unlike structural neuroimaging, fMRI allows for the assessment not only of anatomical integrity but also of functional interactions among brain regions, which is particularly important for understanding systemic disruptions in ASD. The concept of functional connectivity, based on

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 433–446, 2027.
https://doi.org/10.1007/978-3-032-30530-5_27

the synchronization of BOLD time series, enables quantitative evaluation of integrative brain processes and the identification of patterns specific to various neurological and psychiatric conditions. Low-frequency spontaneous BOLD fluctuations have been shown to correlate across functionally connected, including anatomically distant, brain regions, forming large-scale neural networks [2]. Among these networks are the default mode network, the salience network, and the frontoparietal control network, which play key roles in cognitive and socio-emotional processes. In ASD, numerous studies report alterations in functional connectivity, yet findings remain inconsistent, showing both hypo- and hyperconnectivity in various networks, likely due to sample heterogeneity, age-related differences, and methodological factors [3].

To date, a variety of approaches have been proposed for classifying ASD using fMRI data, ranging from classical machine learning methods (logistic regression, support vector machines, random forest) to modern deep learning models, including graph neural networks that incorporate the topology of brain networks [4, 5]. Despite high classification accuracies reported in some studies, the reproducibility of these models remains limited, especially when applied to independent datasets.

A key challenge is the high dimensionality of the data. Representing functional connectivity as a matrix of pairwise interactions between brain regions can result in thousands of potential features even with moderately detailed parcellations. With limited sample sizes, this creates the classical "curse of dimensionality," increasing the risk of overfitting.

Equally important is interpretability. Complex models often operate as "black boxes," making it difficult to determine the contribution of individual connections or networks to the final decision. This limits the clinical utility of such models, as clinicians require not only probabilistic estimates of diagnosis but also insights into the neurobiological basis of these decisions and the identification of structures involved in the pathological process. Moreover, in some studies, reported high accuracies may result from methodological issues such as data leakage between training and test sets, feature selection prior to data splitting, or insufficient control of overfitting.

Another critical aspect is biomarker stability. Functional connectivity analyses can vary substantially depending on data splits, preprocessing parameters, and analytical approaches, even when identical pipelines are applied. In the context of high inter-individual and inter-site variability (e.g., in multicenter datasets such as ABIDE), ensuring reproducibility becomes a key challenge.

Given these issues—the lack of objective biomarkers, heterogeneity of published data, high dimensionality, instability of features, and limited interpretability of models—developing methods for identifying stable, interpretable, and neurobiologically meaningful patterns of functional connectivity is of particular relevance. The present study aims to develop an approach capable of detecting functional connections that are reproducible across different sample splits and exhibit high discriminative power in distinguishing individuals with ASD from neurotypical controls.

2 Material and Methods

2.1 Data Source

The study was carried out on the basis of open data from the ABIDE I (Autistic Brain Imaging Data Exchange I) consortium, a large international project for the exchange of neuroimaging data for the study of autism spectrum disorders (ASD). The project was created to provide the scientific community with open access to standardized fMRI data to accelerate research in the field of autism [6].

Within the framework of this work, the main source of data for analysis is functional MRI recordings at rest, since the purpose of the study is to identify specific areas of the brain that distinguish individuals with ASD from neurotypical ones.

2.2 Data Preprocessing and Sampling

The fMRI data were pre-processed using a standardized CPAC (Configurable Pipeline for the Analysis of Connectomes) pipeline. This pipeline implements typical processing steps, including: time shift and motion correction, normalization, spatial antialiasing, regression of interfering signals, and bandpass temporal filtering. The use of CPAC ensures the comparability of data from different centers and increases the reproducibility of results [7].

After removing records with missing age, sex, and diagnosis values, and excluding subjects with incomplete functional connectivity data, the final sample consisted of 849 subjects (Table 1).

Table 1. Demographic characteristics.

Group	Total	Male	Female
Control	458	373	85
ASD	391	343	48
Total	849	716	133

The gender imbalance in the sample reflects the characteristics of the ABIDE I database formation in the early 2010s, when diagnostic criteria and tools were primarily developed on male cohorts, and ASD in females was systematically underestimated due to symptom masking [8]. This limitation was taken into account when interpreting the results.

2.3 Functional Connectivity Construction

For each study participant, the raw fMRI data consist of BOLD signal time series recorded at each point in the brain. To transition from the voxel to the regional level of analysis, the CC200 atlas [9] was used, which partitions the brain into 200 approximately equal functional regions of interest (ROIs). The CC200 atlas was chosen due to

its functional homogeneity, which is particularly important for the analysis of functional connectivity. Time series were averaged within each region, reducing data dimensionality and minimizing the influence of physiological noise.

Next, a 200 × 200 functional connectivity matrix was computed for each subject. Each element of this matrix reflects the strength of the connection between two regions and was calculated as the Pearson correlation coefficient between the corresponding time series [10]. Correlation values range from –1 to 1: positive values indicate synchronous activity between regions, whereas negative values indicate oppositely directed dynamics, where one region is active while the other is deactivated.

Since the connectivity matrix is symmetric with respect to the main diagonal, and diagonal elements do not convey information about interactions between different regions, only the upper triangle of the matrix was used for analysis. For 200 regions, this results in 19,900 unique connections, each representing a separate feature characterizing the functional interaction between a specific pair of brain regions.

Two covariates, age and sex of the participant, were added to the feature set, as these factors are known to influence measures of functional connectivity [11]. Consequently, each subject was represented by a feature vector of 19,902 elements, suitable for training machine learning models and identifying the most significant brain regions.

2.4 Feature Selection

The original dataset's high dimensionality (19,902 features for 849 subjects; $p/n \approx 23.4$) posed a risk of overfitting, where the model could capture noise instead of true patterns. Conducting a large number of statistical tests without correction also increases false positive results (approximately 995 expected at $\alpha = 0.05$), necessitating multiple comparison correction.

A two-step dimensionality reduction approach combining statistical filtering and feature ranking was implemented. In the first step, a two-sample Welch t-test was applied, followed by Benjamini–Hochberg false discovery rate correction, to identify connections that differed significantly between groups while controlling the proportion of false discoveries. Feature selection was performed exclusively on training data to avoid information leakage and inflated performance estimates.

In the second step, the SelectKBest method with ANOVA F-statistics was used to select the 100 most informative features. This univariate ranking approach evaluates each feature independently for its ability to discriminate between classes. Using ANOVA F-statistics allows quantitative ranking of features by discriminative power, while SelectKBest ensures control over the number of features, important for interpretability and subsequent neurobiological analysis.

Feature selection was performed within cross-validation folds, preventing data leakage and providing an unbiased estimate of model generalizability. Stability was assessed by the frequency of feature occurrence across five folds, resulting in 22 functional connections with the highest reproducibility.

2.5 Classification and Model Evaluation

To solve the classification problem, a multi-stage methodology was developed. First, the dataset was split into training (80%) and test (20%) samples, preserving class proportions. This separation ensures an independent dataset for final evaluation, which is not involved in training or hyperparameter tuning.

A 5-fold cross-validation was applied to the training sample. Within each fold, a two-step feature selection was performed: initially, an FDR-corrected t-test identified statistically significant differences between groups, followed by the SelectKBest method with ANOVA F-statistics to select the 100 most informative connections. All feature selection steps were performed exclusively on training data within each fold to avoid data leakage and ensure unbiased model evaluation. This procedure reduced the feature space from 19,902 to 100 features.

After selection, features were standardized, and a logistic regression model with L2 regularization was trained. Logistic regression was chosen for its interpretability: the sign of each coefficient directly indicates the direction of change in functional connectivity (hyper- or hypoconnectivity), allowing straightforward neurobiological interpretation. Unlike more complex models, logistic regression provides a balance between classification performance and interpretability.

The stability of selected features was assessed by their frequency across the five cross-validation folds. Features consistently selected in each fold were considered the most reliable potential biomarkers. The final model was trained on the entire training set and evaluated on the held-out test set to assess generalizability.

Model performance was evaluated using several metrics to capture different aspects of classification quality. ROC-AUC was the primary metric because it reflects the model's ability to distinguish between classes independently of the threshold and allows comparison with other studies. Sensitivity was used to assess the proportion of correctly classified ASD participants, which is critical for diagnostic purposes, as missing a case has more serious consequences than a false positive. Specificity reflects the proportion of correctly classified neurotypical controls. Accuracy is reported for comparison with other studies but is not the primary metric for unbalanced datasets (Table 2).

Table 2. Logistic regression model performance on training and test samples.

Sample	Accuracy	AUC	Sensitivity	Specificity
Train	0.7202	0.8137	0.7668	0.680
Test	0.7176	0.8052	0.7436	0.6957

The minimal difference between training and test performance indicates that the model generalizes well and is not overfitted. Of particular importance is the high sensitivity: the model correctly identifies three out of four ASD cases, which is critical for potential clinical applications.

The ROC curves for the training and test sets, along with the confusion matrix, are presented in Fig. 1. This figure visualizes both the discriminatory power of the model

across thresholds and the classification outcomes, highlighting the reliability of predicted labels.

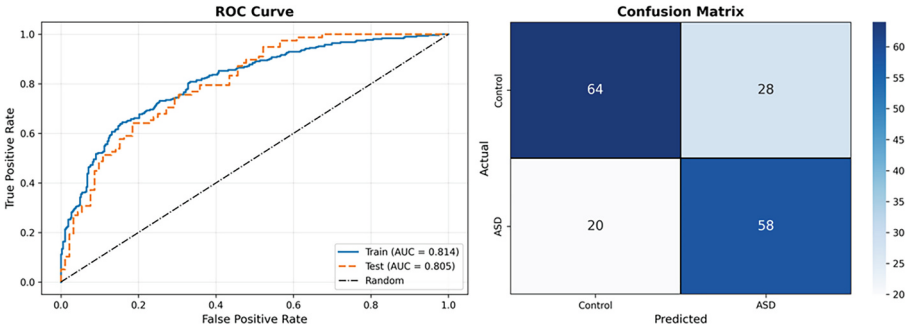


Fig. 1. ROC curves and confusion matrix for the logistic regression model.

To evaluate whether nonlinear decision boundaries could improve classification performance, an additional experiment was conducted using a Support Vector Machine (SVM) with a radial basis function (RBF) kernel. The same data partitioning strategy, preprocessing pipeline, and feature selection procedure were applied. Hyperparameters were optimized on the training data using grid search.

The nonlinear SVM achieved performance comparable to logistic regression on the independent test set, without providing a substantial improvement in classification performance (Table 3).

Table 3. SVM-RBF model performance on training and test samples.

Sample	Accuracy	AUC	Sensitivity	Specificity
Train	0.7275	0.8184	0.7923	0.6721
Test	0.7176	0.8060	0.7692	0.6739

Although the nonlinear SVM slightly improved sensitivity, the overall classification performance remained nearly identical to logistic regression. Given the substantially higher interpretability of logistic regression coefficients for neurobiological interpretation, logistic regression was retained as the primary model for subsequent interpretation of stable functional connections.

2.6 Decoding and Mapping Stable Functional Connections

After identifying 22 stable functional connections consistently selected across all five cross-validation folds, the next step was their anatomical identification and mapping to functional brain networks for neurobiological interpretation. The original features, represented as FC_N indices, were transformed into specific pairs of brain regions. Each

feature corresponds to a unique pair of ROIs located in the upper triangle of a 200 × 200 functional connectivity matrix. The linear index N was converted to ROI numbers (i, j), which were then mapped to anatomical names and coordinates from the CC200 atlas in standard MNI space. For each connection, both regions involved and the direction of change (hyper- or hypoconnectivity) were determined from the sign of the logistic regression coefficient.

For system-level interpretation, the 28 unique ROIs involved in the 22 stable connections were mapped to the Yeo7 functional networks [12], which divide the cerebral cortex into seven large-scale networks: Visual, Somatomotor, Dorsal Attention, Ventral Attention, Limbic, Frontoparietal, and Default. Yeo7 mapping was chosen because the brain functions as an integrated system, and it allows moving from isolated anatomical regions to network-level interpretation, enabling identification of functional systems affected in ASD and comparison with contemporary network-based studies. Alternative approaches, such as mapping to Brodmann cytoarchitectonic areas, are technically more challenging, requiring additional transformations, whereas Yeo7 mapping based on MNI coordinates provides a direct and methodologically sound approach.

The mapping was performed using ROI centroids in MNI space extracted from the CC200 atlas and the nilearn library, which identifies the nearest functional network based on Euclidean distance. To improve reliability for ROIs located near network boundaries, an extended search radius of 10 mm was applied. Subcortical structures not included in Yeo7 cortical networks were assigned to a Background category for separate analysis (Fig. 2).

Table 4. Distribution of ROIs and stable functional connections across Yeo7 networks.

Yeo7 Network	Number of ROIs	Total Connections	Enhanced in ASD	Weakened in ASD
Default	12	21	4	17
Somatomotor	6	7	3	4
Background	5	9	6	3
Frontoparietal	2	2	1	1
Ventral Attention	2	3	2	1
Visual	1	2	2	0

This decoding and mapping procedure allowed moving from abstract indices to concrete anatomical and functional structures, providing a basis for interpreting altered functional connectivity in ASD within large-scale neural networks.

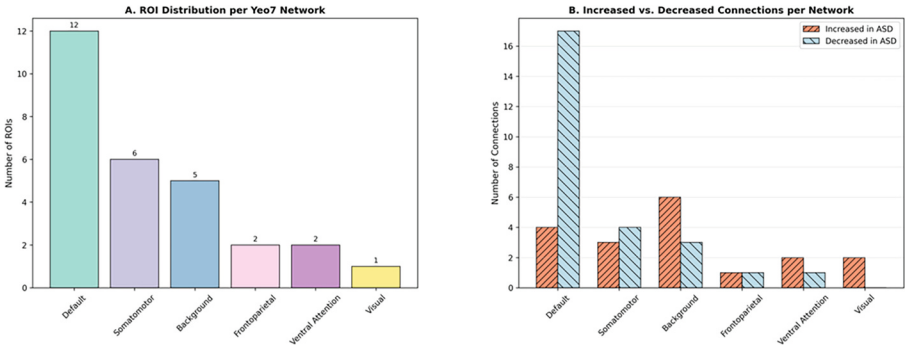


Fig. 2. Distribution of stable functional connections across Yeo7 networks. (A) Number of ROIs per network. (B) Number of connections increased vs. decreased in ASD per network.

3 Discussion

The obtained classification accuracy is comparable to the results of modern studies using the ABIDE dataset. A comparison of key performance metric is presented in Table 5.

Table 5. Accuracy comparison with studies using rs-fMRI data

Research	Accuracy
Abraham et al. (2017) [13]	67%
Heinsfeld et al. (2018) [4]	70%
Dong et al. (2025) [14]	72%
Our study	71.8%

The studies presented in Table 5 are based on different machine learning approaches, including deep neural networks and graph-based models. Despite relying on a comparatively simple and interpretable logistic regression framework, the proposed methodology achieved classification accuracy comparable to these approaches, while additionally allowing direct neurobiological interpretation and identification of stable functional biomarkers, which is particularly important for potential clinical applications.

Among the 100 selected features, 22 functional connections were consistently identified across all five cross-validation folds. Their repeated selection across different training subsets suggests that these connections are relatively stable and not strongly dependent on a particular data split. The complete list of reproducible ROI-to-ROI functional connections is presented in Appendix (Table 6). It should be noted that some anatomical labels appear multiple times because the CC200 atlas contains several distinct ROIs localized within the same anatomical structure but belonging to different functional regions.

The additional comparison with a nonlinear SVM classifier demonstrated performance nearly identical to logistic regression on the independent test set. This may indicate that, after feature selection, the differences between ASD and control groups can already be effectively captured by a linear model. Importantly, increasing model complexity did not lead to a substantial improvement in classification performance, which may suggest limited effectiveness of nonlinear models for rs-fMRI classification tasks after feature space reduction.

At the same time, logistic regression offers an important practical advantage in terms of interpretability. Unlike nonlinear "black-box" models, logistic regression coefficients can be directly interpreted in terms of increases or decreases in functional connectivity. This makes it possible to relate classification results to specific neurobiological mechanisms and identify reproducible biomarkers that may be clinically meaningful.

An important limitation of the present study is the pronounced gender imbalance of the ABIDE I dataset, where males constitute the majority of participants. This imbalance is partly related to the fact that earlier ASD research primarily focused on male cohorts, while autism in females was diagnosed less frequently [8]. Consequently, the 22 stable functional connections identified in this work may predominantly reflect connectivity patterns characteristic of the male ASD phenotype. Previous neuroimaging studies suggest that females with ASD may show some differences in brain organization and functional connectivity patterns, including regions associated with social cognition, the Default Mode Network (DMN), and mechanisms involved in symptom compensation [15]. Therefore, additional validation in larger and more sex-balanced cohorts is necessary before generalizing the identified biomarkers to female ASD populations.

Mapping of stable functional connections onto the Yeo7 network revealed a characteristic pattern of alterations in ASD: pronounced hypoconnectivity within the Default Mode Network (17↓ / 4↑), accompanied by relative hyperconnectivity in subcortical regions (6↑ / 3↓) and the visual network (2↑). Overall, these results are consistent with current views of ASD as a disorder involving altered organization of large-scale brain networks.

The observed hypoconnectivity of the DMN is supported by a recent large-scale meta-analysis [16], which integrated data from functional and structural neuroimaging studies. The authors reported reduced functional activity in core DMN regions, including the left angular gyrus, bilateral anterior cingulate cortex, and medial prefrontal cortex. Several of these regions, including the angular gyrus and anterior cingulate cortex, were also represented among the stable ROIs identified in our study, supporting the reproducibility of DMN-related alterations. Importantly, the meta-analysis revealed partial overlap between functional and structural abnormalities across multiple brain regions, suggesting that ASD is associated with both functional and anatomical alterations.

The increased involvement of subcortical structures observed in our results is consistent with findings from Hadjikhani et al. [17], who demonstrated abnormally elevated activation in subcortical regions—including the superior colliculus, pulvinar nucleus of the thalamus, and amygdala—during constrained eye contact in individuals with ASD. In our study, the thalamus and midbrain show a predominance of hyperconnectivity (6↑ / 3↓), which may reflect altered sensory gating and atypical processing of socially relevant stimuli. However, given that [17] is based on task-related fMRI, this correspondence

should be interpreted as indirect rather than direct evidence of resting-state connectivity alterations.

The mixed pattern observed in the somatomotor network (3↑ / 4↓) suggests multi-directional alterations of functional connectivity in ASD. In particular, the postcentral gyrus and cerebellum in our results participate in both increased and decreased connections, indicating that connectivity changes within the sensorimotor system do not follow a single consistent pattern. These findings are consistent with previous work [18], which reported reduced connectivity within the sensorimotor network together with increased interactions between the sensorimotor and other functional networks in ASD. This may reflect atypical processing and coordination of sensorimotor information rather than only alterations in motor function.

The observed hyperconnectivity within the visual network may reflect atypical weighting of perceptual processing in ASD. This interpretation is consistent with theories emphasizing enhanced or preferential processing of visual information; however, it should be considered a hypothesis requiring further validation in studies directly linking functional connectivity patterns with behavioral measures.

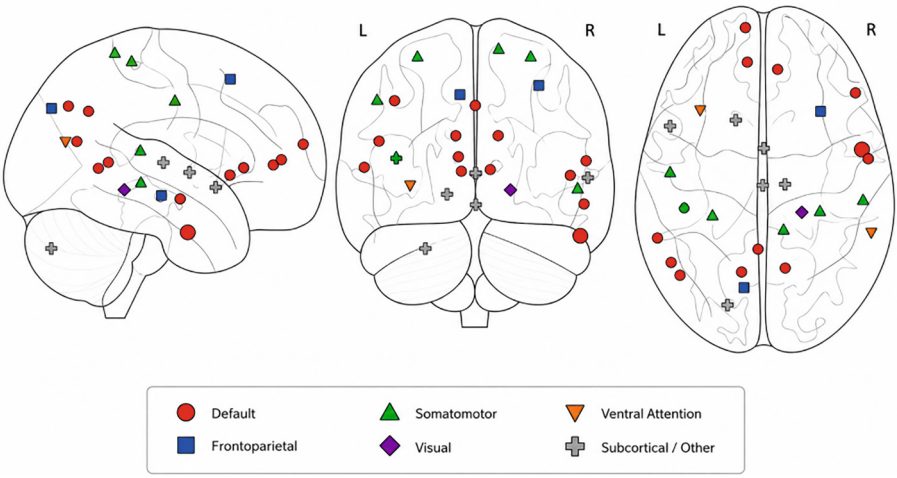


Fig. 3. Localization of reproducible brain regions in ASD by Yeo7 network parcellation.

As shown in Fig. 3, the 22 stable ROIs identified in ASD are distributed across multiple Yeo7 networks. While the figure illustrates the spatial distribution, the predominance of DMN hypoconnectivity and the involvement of subcortical and visual systems are quantitatively summarized in Table 4 and supported by the direction of logistic regression coefficients.

4 Conclusion

In this study, a methodology was developed to identify stable functional connections that distinguish individuals with autism spectrum disorder (ASD) from neurotypical controls using resting-state functional MRI data. Open-access data from the ABIDE consortium

were analyzed, including 849 subjects and 19,902 functional connections calculated using the CC200 atlas.

To address the high dimensionality of the data ($p/n \approx 23.4$) and control for multiple comparisons, a two-stage feature selection approach was applied: first, the Welch t-test with FDR correction, followed by the SelectKBest method with ANOVA F-statistics to select the 100 most informative connections. A critical methodological decision was to perform feature selection exclusively on the training data within each cross-validation fold, thereby preventing data leakage and providing an unbiased assessment of model generalizability. The stability of feature selection was evaluated by the frequency of feature occurrence across all five folds, which made it possible to identify 22 functional connections with the highest reliability.

On the independent test set, the final logistic regression model achieved an accuracy of 71.8%, an AUC of 80.5%, a sensitivity of 74.4%, and a specificity of 69.6%. The small difference between training and test performance indicates limited overfitting and good generalizability of the model. A comparison with the nonlinear SVM-RBF classifier demonstrated nearly identical results, suggesting that more complex nonlinear approaches do not provide a substantial advantage within the reduced feature space used in this study. At the same time, the obtained performance is comparable to that reported in recent ABIDE-based studies, including approaches using neural network architectures, supporting the validity of the selected interpretable model.

Mapping the identified ROIs onto the Yeo7 functional networks revealed that the detected alterations were predominantly localized within the Default Mode Network and were mainly characterized by hypoconnectivity. These findings are consistent with large meta-analyses and further support the central role of the DMN in ASD-related social dysfunction. In addition, relative hyperconnectivity of subcortical structures, including the thalamus and midbrain, was observed and may reflect altered sensory gating and atypical processing of socially relevant stimuli. Hyperconnectivity within the visual network may indicate atypical weighting of perceptual processing, although this interpretation requires further validation. The somatomotor network demonstrated a mixed pattern of increased and decreased connectivity, likely reflecting the heterogeneity of motor and sensory manifestations associated with ASD.

The patterns identified in this study are consistent with findings reported in previous research, supporting the reproducibility of the selected biomarkers and their correspondence with current neurobiological models of ASD. At the same time, several limitations should be considered. The ABIDE I dataset is characterized by a substantial gender imbalance, reflecting diagnostic practices of the early 2010s and the under-recognition of ASD in females due to symptom masking and features of the female ASD phenotype. As a result, the detected biomarkers may primarily reflect connectivity patterns characteristic of male ASD populations. Additional validation is therefore required in larger and more balanced samples, as well as in ethnically diverse and independent cohorts.

Further studies may examine whether the observed functional connectivity alterations are associated with the severity of ASD symptoms measured using clinical scales such as ADOS and ADI-R. This may help clarify the clinical relevance of the proposed biomarkers. In addition, alterations observed in visual and subcortical systems may be related to atypical sensory processing in ASD and could represent potential targets for

new therapeutic approaches. However, these assumptions require additional validation in clinical settings.

Thus, the proposed approach for detecting stable ASD biomarkers represents a step toward more objective diagnostic methods and a better understanding of the neurobiological mechanisms underlying the disorder. The 22 reproducible functional connections identified in this study may serve as a foundation for the development of medical decision-support systems and for further research into the pathogenesis of autism spectrum disorder.

Appendix

Table 6. Stable functional connections identified across all cross-validation folds

ROI 1	ROI 2	Connectivity in ASD
Precuneus (19)	Middle temporal gyrus (140)	Decreased
Middle temporal gyrus (140)	Precuneus (174)	Decreased
Anterior cingulate (5)	Middle temporal gyrus (140)	Decreased
Precuneus (3)	Middle temporal gyrus (140)	Decreased
Midbrain (68)	Middle frontal gyrus (127)	Decreased
Midbrain (68)	Superior temporal gyrus (153)	Increased
Insula (20)	Inferior frontal gyrus (144)	Decreased
Cerebellum (120)	Transverse temporal gyrus (200)	Increased
Superior parietal lobule (65)	Precuneus (147)	Increased
Angular gyrus (2)	Middle temporal gyrus (140)	Decreased
Postcentral gyrus (96)	Postcentral gyrus (154)	Decreased
Thalamus (18)	Middle temporal gyrus (107)	Increased
Anterior cingulate (22)	Middle temporal gyrus (140)	Decreased
Thalamus (18)	Middle temporal gyrus (117)	Increased
Middle temporal gyrus (69)	Thalamus (148)	Increased
Frontal pole (104)	Middle temporal gyrus (140)	Decreased
Parahippocampal gyrus (48)	Middle temporal gyrus (69)	Increased
Thalamus (18)	Caudate (47)	Decreased
Postcentral gyrus (28)	Postcentral gyrus (154)	Decreased
Parahippocampal gyrus (48)	Middle temporal gyrus (117)	Increased
Angular gyrus (82)	Middle temporal gyrus (140)	Decreased
Midbrain (68)	Middle temporal gyrus (117)	Increased

References

1. National Center for Biotechnology Information (US): Autism spectrum disorder. In: StatPearls. StatPearls Publishing, Treasure Island (FL) (2023). https://www.ncbi.nlm.nih.gov/books/NBK525976/. Accessed 28 Mar 2026
2. Biswal, B., Yetkin, F.Z., Haughton, V.M., Hyde, J.S.: Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn. Reson. Med. **34**(4), 537–541 (1995). https://doi.org/10.1002/mrm.1910340409
3. Uddin, L.Q., Supekar, K., Menon, V.: Reconceptualizing functional brain connectivity in autism from a developmental perspective. Front. Hum. Neurosci. **7**, 458 (2013). https://doi.org/10.3389/fnhum.2013.00458
4. Heinsfeld, A.S., Franco, A.R., Craddock, R.C., Buchweitz, A., Meneguzzi, F.: Identification of autism spectrum disorder using deep learning and the ABIDE dataset. Neuroimage Clin. **17**, 16–23 (2018). https://doi.org/10.1016/j.nicl.2017.08.017
5. Parisot, S., et al.: Disease prediction using graph convolutional networks: application to autism spectrum disorder and Alzheimer's disease. Med. Image Anal. **48**, 117–130 (2018). https://doi.org/10.1016/j.media.2018.06.001
6. Di Martino, A., Yan, C.G., Li, Q., Denio, E., Castellanos, F.X., Alaerts, K., et al.: The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol. Psychiatry **19**(6), 659–667 (2014). https://doi.org/10.1038/mp.2013.78
7. Craddock, R.C., et al.: Towards automated analysis of connectomes: the configurable pipeline for the analysis of connectomes (C-PAC). In: Frontiers in Neuroinformatics Conference Abstract (2013). https://doi.org/10.3389/conf.fninf.2013.09.00042
8. Hull, L., Petrides, K.V., Mandy, W.: The female autism phenotype and camouflaging: a narrative review. Rev. J. Autism Dev. Disord. **7**(4), 306–317 (2020). https://doi.org/10.1007/s40489-020-00197-9
9. Craddock, R.C., James, G.A., Holtzheimer, P.E., Hu, X.P., Mayberg, H.S.: A whole brain fMRI atlas generated via spatially constrained spectral clustering. Hum. Brain Mapp. **33**(8), 1914–1928 (2012). https://doi.org/10.1002/hbm.21333
10. Jin, H., et al.: Is Pearson's correlation coefficient enough for functional connectivity in fMRI? Imaging Neurosci. **3**, imag_a_1052 (2025). https://doi.org/10.1162/IMAG.a.1052
11. Zhang, C., Cahill, N.D., Arbabshirani, M.R., White, T., Baum, S.A., Michael, A.M.: Sex and age effects of functional connectivity in early adulthood. Brain Connect. **6**(9), 700–713 (2016). https://doi.org/10.1089/brain.2016.0429
12. Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., et al.: The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J. Neurophysiol. **106**(3), 1125–1165 (2011). https://doi.org/10.1152/jn.00338.2011
13. Abraham, A., et al.: Deriving reproducible biomarkers from multi-site resting-state data: an autism-based example. Neuroimage **147**, 736–745 (2017). https://doi.org/10.1016/j.neuroimage.2016.10.045
14. Dong, Y., Batalle, D., Deprez, M.: A framework for comparison and interpretation of machine learning classifiers to predict autism on the ABIDE dataset. Hum. Brain Mapp. **46**(5), e70190 (2025). https://doi.org/10.1002/hbm.70190
15. Walsh, M.J., Baxter, L.C., Smith, C.J., Braden, B.B.: Age and sex differences in autism spectrum disorder: a systematic review of brain-based sex differences across the lifespan. Front. Psychiatry **12**, 689303 (2021). https://doi.org/10.3389/fpsyt.2021.689303
16. Guo, Z., Tang, X., Xiao, S., Yan, H., Sun, S., Yang, Z., et al.: Systematic review and meta-analysis: multimodal functional and anatomical neural alterations in autism spectrum disorder. Mol. Autism **15**, 16 (2024). https://doi.org/10.1186/s13229-024-00593-6

17. Hadjikhani, N., et al.: Look me in the eyes: constraining gaze in the eye-region provokes abnormally high subcortical activation in autism. Sci. Rep. **7**, 3163 (2017). https://doi.org/10.1038/s41598-017-03378-5
18. Minnigulova, A., Dragoy, O., Arutiunian, V.: Atypical segregation of frontoparietal and sensorimotor networks is related to social and executive function impairments in children with ASD. Brain Imaging Behav. **19**(4), 846–855 (2025). https://doi.org/10.1007/s11682-025-01016-7

TFPACO: Traffic Flow Planning System for Software-Defined Networks Based on Ant Colony Optimization Algorithm

Aleksandr Safonenko and Vladimir Korkhov(✉)

Saint Petersburg State University, 7-9 Universitetskaya Embankment, St. Petersburg 199034, Russia
st107536@student.spbu.ru, v.korkhov@spbu.ru

Abstract. The rapid growth of the Internet due to the COVID pandemic and the increasing need for video content are creating pressure on existing network infrastructure. Therefore, there is a need for efficient Quality of Service routing techniques. In this context, this paper presents a new algorithm for Traffic Flow Planning system in SDN based on Ant Colony Optimization (TFPACO). In this algorithm, unlike existing optimization techniques, mice and elephant flows are considered separately. Mice flows are optimized using a Dijkstra algorithm considering risk and delay while elephant flows use a multi-objective ant colony optimization model. As a result we get multiple solutions, and only nondominated solutions are stored in the archive. Among the solutions obtained in the archive, a set of solutions is chosen for proportional routing. In this context, Dijkstra's algorithm, multi-objective Dijkstra's algorithm, TSACO, and the Motlagh algorithm are used as reference algorithms. A discrete-time simulation environment with dynamic edge parameters is used to check the performance of the proposed algorithm. The simulation results show that the proposed algorithm entirely eliminates flow and demand blocking with less latency than existing algorithms.

Keywords: software-defined networking · ant colony optimization · multi-objective optimization · flow scheduling · quality of service

1 Introduction

Software-defined networking (SDN) has become one of the main architectural approaches for improving flexibility and controllability of modern communication networks. By separating the control plane from the data plane, SDN enables centralized management of routing decisions and provides a global view of the current network state [1,2]. This property makes SDN particularly suitable for traffic engineering and traffic management tasks, where route selection must adapt to varying load conditions and heterogeneous application requirements.

However, with the increasing volume of network traffic, there are many challenges for SDN traffic management. First, route selection is no longer a simple

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 447–463, 2027.
https://doi.org/10.1007/978-3-032-30530-5_28

shortest path problem. In practice, route selection should be based on various network factors, including delay, reliability, and bandwidth [3]. Second, heterogeneous traffic flows cause conflicting quality-of-service (QoS) requirements. In particular, mice flows are usually short and delay-sensitive, whereas elephant flows require a lot of available bandwidth and may create local congestion if routed improperly [4,5]. Therefore, a good SDN traffic management algorithm should be able to handle the differences between various types of traffic and multiple quality-of-service requirements.

A large number of SDN traffic management methods have been proposed in recent years. Existing researches include single-objective delay-based routing, weighted multi-criteria routing, load-balancing methods, and routing algorithms based on swarm intelligence [2,4,6]. The classical Dijkstra algorithm is simple and computationally efficient. However, this algorithm does not take into account the remaining bandwidth and reliability. The multi-objective Dijkstra algorithm improves flexibility by combining several metrics into a single scalar weight [6]. Nevertheless, the results are heavily dependent on the weights used. In the TSACO algorithm, mice traffic and elephant traffic are managed separately using ant colony optimization and edge disjoint multipath [5]. The multi-objective ACO algorithm proposed by Motlagh et al. focuses on route construction with risk and cost criteria and produces nondominated solutions [7].

Other related studies in the context QoS routing in SDN of Internet of Things based on swarm intelligence have also been conducted. For example, bio-inspired routing based on ACO has been used to analyze delay, loss, jitter, and rule capacity constraints in QoS-aware routing [8]. These studies demonstrate the potential of ant colony optimization in traffic flow planning under varying network requirements.

In this paper, a new Traffic Flow Planning system based on Ant Colony Optimization (TFPACO) algorithm is proposed. The method combines Dijkstra algorithm for mice flows with multi-objective ACO algorithm for elephant flows. The candidate paths for elephant flows are determined by path delay, path risk, and path inverse bottleneck remaining bandwidth. The nondominated paths are then stored in the archive and then used for multipath traffic allocation. Our proposed algorithm is compared to classical Dijkstra, multi-objective Dijkstra, TSACO, and the Motlagh algorithms.

The rest of the paper is outlined as follows. Section 2 presents the problem statement and mathematical model. Section 3 overviews the baseline routing algorithms. Section 4 overviews the proposed TFPACO algorithm. Section 5 presents the simulation model and experimental results. Section 6 overviews the obtained results and concludes the paper.

2 Problem Statement and Mathematical Model

The considered SDN topology is represented as an undirected graph

$$G = (V, E), \tag{1}$$

where V is the set of nodes and E is the set of edges. Each edge $(i, j) \in E$ is associated with the following parameters:

$$d_{ij} > 0, \qquad b_{ij} > 0, \qquad p_{ij} \in (0, 1), \tag{2}$$

where d_{ij} is link delay, b_{ij} is link bandwidth, and p_{ij} is the packet loss probability of the link.

A traffic flow is represented as

$$q = (s, s', \lambda, \xi, c), \tag{3}$$

where $s \in V$ is the source node, $s' \in V$ is the destination node, $\lambda > 0$ is the bandwidth demand, ξ is the flow lifetime, and $c \in \{\text{mice}, \text{elephant}\}$ is the traffic class.

A route from s_q to s'_q is represented as a path

$$\mathcal{P}_q = (v_0, v_1, \ldots, v_\ell), \tag{4}$$

where $v_0 = s_q$, $v_\ell = s'_q$, and $(v_r, v_{r+1}) \in E$ for all $r = 0, \ldots, \ell - 1$. In the optimization model, the same route may be represented as binary edge variables

$$x_{ij,q} = \begin{cases} 1, & \text{if edge } (i, j) \text{ belongs to } \mathcal{P}_q, \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

For elephant flows, multipath routing is allowed. In that case, the controller selects a set of routes

$$\{\mathcal{P}_q^{(1)}, \mathcal{P}_q^{(2)}, \ldots, \mathcal{P}_q^{(m_q)}\} \tag{6}$$

together with traffic shares

$$\boldsymbol{y}_q = (y_q^{(1)}, y_q^{(2)}, \ldots, y_q^{(m_q)}), \tag{7}$$

such that

$$y_q^{(r)} \geq 0, \qquad \sum_{r=1}^{m_q} y_q^{(r)} = 1. \tag{8}$$

Hence, the amount of traffic sent over route $\mathcal{P}_q^{(r)}$ is equal to $y_q^{(r)} \lambda_q$.

Let $\mathcal{F}(t)$ be the set of active flows at time t, and let

$$b_{ij}^{\text{used}}(t) = \sum_{q \in \mathcal{F}(t)} \sum_{r=1}^{m_q} \lambda_q y_q^{(r)} x_{ij,q}^{(r)}, \tag{9}$$

where $m_q = 1$ and $y_q^{(r)} = 1$, if $c_q = \text{mice}$, be the used bandwidth of edge (i, j) at time t. Then the remaining bandwidth is

$$b_{ij}^{\text{rem}}(t) = b_{ij} - b_{ij}^{\text{used}}(t). \tag{10}$$

For a feasible route, the remaining bandwidth must be nonnegative on all its edges, and for all active flows the bandwidth constraints must satisfy

$$b_{ij}^{\text{used}} \leq b_{ij}, \qquad \forall (i,j) \in E. \tag{11}$$

The route delay is assumed additive [5]:

$$f_1(\mathcal{P}_q) = D(\mathcal{P}_q) = \sum_{(i,j)\in\mathcal{P}_q} d_{ij}. \tag{12}$$

This is the standard path delay model used in shortest-path and QoS routing.

The route risk is also assumed to be additive:

$$f_2(\mathcal{P}_q) = R(\mathcal{P}_q) = \sum_{(i,j)\in\mathcal{P}_q} -\ln(1-p_{ij}), \tag{13}$$

which is consistent with the path risk formulation used in [7]. This representation is convenient because the packet delivery probability of the path is given as

$$\text{PDR}(\mathcal{P}_q) = \prod_{(i,j)\in\mathcal{P}_q} (1-p_{ij}) = \exp(-R(\mathcal{P}_q)). \tag{14}$$

For elephant flows, an additional QoS criteria is the bottleneck remaining bandwidth [5],

$$B(\mathcal{P}_q) = \min_{(i,j)\in\mathcal{P}_q} b_{ij}^{\text{rem}}, \tag{15}$$

which characterizes the narrowest edge on the route. Since all optimization criteria are considered in minimization form, the bandwidth-related objective is defined as

$$f_3(\mathcal{P}_q) = \frac{1}{B(\mathcal{P}_q)}. \tag{16}$$

Thus, routes with larger bottleneck remaining bandwidth are preferred.

For a single route, the objective vector is defined as

$$F(\mathcal{P}_q) = \left(f_1(\mathcal{P}_q),\, f_2(\mathcal{P}_q),\, f_3(\mathcal{P}_q)\right). \tag{17}$$

The routing problem is therefore formulated as a multi-objective optimization problem

$$\min_{\mathcal{P}_q\in S_q} \left\{F(\mathcal{P}_q)\right\}, \tag{18}$$

where S_q is the set of feasible routes for flow q.

The aim is to find feasible routes, or sets of routes, that provide a good compromise between delay, reliability, and remaining bandwidth.

A route $\mathcal{P}_q^*$ is Pareto-optimal if there exists no other feasible route $\mathcal{P}_q$ such that

$$f_i(\mathcal{P}_q) \leq f_i(\mathcal{P}_q^*), \qquad \forall i, \tag{19}$$

and for at least one index j

$$f_j(\mathcal{P}_q) < f_j(\mathcal{P}_q^*). \tag{20}$$

In other words, a Pareto-optimal route cannot be improved in one objective without worsening at least one other objective.

In the proposed algorithm, mice and elephant flows are treated differently. Mice flows are routed over a single path due to their small size and high delay sensitivity. Elephant flows, in contrast, may be distributed over multiple routes. For the selected elephant route with bottleneck remaining bandwidth values $B_1, \ldots, B_{m_q}$, the traffic shares are assigned proportionally [5]:

$$y_q^{(r)} = \frac{B_r}{\sum_{j=1}^{m_q} B_j}, \qquad \sum_{r=1}^{m_q} y_q^{(r)} = 1. \tag{21}$$

This allocation rule makes the routing decision bandwidth-aware and allows the centralized controller to distribute demand of elephant flow across multiple feasible paths.

Therefore, the problem considered in this work is to construct feasible routes in SDN under multiple QoS criteria and choose the best of them.

3 Overview of Routing Methods

This section briefly overviews the routing algorithms used as baselines in the comparative study. The considered algorithms differ in the way they handle QoS criteria, mice and elephant flows, and route construction under load.

3.1 Classical Dijkstra Algorithm

The classical Dijkstra algorithm is the standard shortest-path algorithm for graphs with nonnegative edge weights. In the SDN the edge weight is chosen as the link delay:

$$w_{ij} = d_{ij}. \tag{22}$$

Then the algorithm returns a minimum-delay route between the source and the destination nodes.

The main advantages of this approach are low computational complexity and simple implementation. However, the algorithm is single-objective: it ignores remaining bandwidth, packet loss, and route diversity. As a result, the delay-based shortest path may become overloaded under high traffic demand, especially in the presence of elephant flows.

3.2 Multi-objective Dijkstra Algorithm

The multi-objective Dijkstra algorithm considered in this work follows the weighted-sum idea used in [6]. In that approach, several edge parameters are combined into a single edge cost:

$$w_{ij} = \omega_1 \widetilde{d}_{ij} + \omega_2 \widetilde{b}_{ij}^{\mathrm{rem}} + \omega_3, \tag{23}$$

where $\widetilde{d}_{ij}$ is the normalized delay, $\widetilde{b}_{ij}^{\text{rem}}$ is the normalized remaining bandwidth, and the constant term is the hop-count contribution.

Hop count is a characteristic of a path defined as the total number of traversed edges. In the algorithm each edge contributes one unit to the total hop count. As a result, this unit contribution along the path represents the total number of hops. Therefore, in the edge-weight representation, the hop-count term is written as a constant value for each edge.

In the paper [6], three weight configurations are considered:

$$\{1, -0.1, 0.2\}, \qquad \{1, -0.5, 0.2\}, \qquad \{0.5, -1, 0.2\}.$$

Compared with the classical Dijkstra algorithm, the multi-objective one is more flexible and better takes into account QoS criteria. However, all considered criteria are still reduced to a single scalar weight, and the result remains sensitive to the selected coefficients.

3.3 TSACO Algorithm

The TSACO method proposed in [5] is specifically designed for SDN traffic scheduling under mice and elephant flow coexistence. Its main idea is to handle these traffic classes separately. Mice flows are delay-sensitive and are routed using a delay-oriented ACO algorithm, while elephant flows are bandwidth-sensitive and are routed by a bandwidth-oriented ACO algorithm.

The probability of choosing the next edge is defined by the standard ACO transition rule:

$$P_{ij} = \begin{cases} \dfrac{\tau_{ij}^{\alpha}\eta_{ij}^{\beta}}{\sum_{k\in\text{neighbours}_i}\tau_{ik}^{\alpha}\eta_{ik}^{\beta}}, & \text{if } j \in \text{neighbours}_i, \\ 0, & \text{otherwise}, \end{cases} \tag{24}$$

where τ_{ij} is the pheromone level and η_{ij} is the heuristic desirability of edge (i, j). In TSACO, the heuristic differs for the two classes of flows. For mice flows, the heuristic is delay-oriented:

$$\eta_{ij} = \frac{1}{d_{ij}}, \tag{25}$$

whereas for elephant flows the heuristic is based on remaining bandwidth:

$$\eta_{ij} = b_{ij}^{\text{rem}}. \tag{26}$$

For elephant traffic, TSACO applies a k-path ACO algorithm. Several edge-disjoint routes are constructed iteratively until the total available bandwidth becomes sufficient. The flow demand is then distributed proportionally to route bandwidth. This makes TSACO an important reference for SDN traffic scheduling, which takes into account different flow classes. At the same time, the method is focused primarily on bandwidth efficiency and does not explicitly model route risk as an optimization objective.

3.4 Motlagh Multi-objective ACO Algorithm

The routing algorithm proposed by Motlagh et al. [7] formulates SDN route restoration as a multi-objective optimization problem under two criteria: path risk and operational cost. In that method, the path risk is defined as

$$\beta(\mathcal{P}_q) = \sum_{(i,j)\in\mathcal{P}_q} -\ln(1-p_{ij}), \tag{27}$$

where p_{ij} is the packet loss probability. The operational cost reflects the route reconfiguration overhead when replacing an already installed route by a new one.

The algorithm combines ACO route construction with mechanisms taken from multi-objective evolutionary optimization. At each iteration, candidate routes are generated by ants, evaluated with respect to the two objectives, and filtered by nondominated sorting. When the number of retained solutions exceeds the desired archive size, crowding distance is used to preserve diversity among the nondominated routes. The pheromone matrix is then updated using the quality of the obtained solutions.

An additional feature of the Motlagh algorithm is the penalization of the last edge of a failed route. This discourages repeated exploration of unproductive directions and improves the exploration ability of the ACO. After the final nondominated set is obtained, a compromise route is chosen by knee-point analysis.

The Motlagh algorithm is particularly relevant to the present work because it provides a clear example of multi-objective ACO routing in SDN. However, its original formulation focuses on route restoration and does not explicitly address class-aware traffic planning or bottleneck bandwidth-driven multipath allocation.

3.5 Discussion

The considered baseline methods illustrate different approaches to SDN routing. Classical Dijkstra is computationally efficient but only delay-oriented. Multi-objective Dijkstra improves metric awareness, but still relies on scalarization. TSACO introduces class-aware routing and multipath allocation for elephant flows, but it only takes into account one criteria for each class. The Motlagh algorithm is multi-objective and archive-based, but it does not address mice and elephant flows separately. These observations motivate the design of the proposed TFPACO algorithm, which combines mice and elephant flows with multi-objective ACO algorithm in the SDN.

4 Proposed TFPACO Algorithm

The proposed Traffic Flow Planning Ant Colony Optimization (TFPACO) algorithm combines Dijkstra algorithm for mice flows with multi-objective ACO algorithm for elephant flows. The main idea is to use a lightweight routing algorithm for short delay-sensitive flows while applying a more flexible multi-objective routing algorithm for large bandwidth-sensitive flows.

4.1 General Structure of the Method

For each incoming flow, the centralized controller first determines its traffic class. Mice flows are routed over a single path, whereas elephant flows may be distributed over several routes. This separation reflects the fact that the two traffic classes have different QoS requirements and different impact on network load.

4.2 Mice-Flow Routing

Mice flows are routed using the Dijkstra algorithm based on edge delay and risk. For edge $(i, j) \in E$, the routing weight is defined as

$$w_{ij} = d_{ij} + \phi r_{ij}, \tag{28}$$

where

$$r_{ij} = -\ln(1 - p_{ij}), \tag{29}$$

and ϕ is the coefficient controlling the influence of risk.

For a mice flow, the controller computes a single feasible shortest path with respect to this edge weight. This formulation favors routes with low delay and low risk, while preserving the computational efficiency of Dijkstra-based routing. Since mice flows have relatively small demand, a single-path decision is sufficient in most cases and helps to avoid unnecessary control overhead.

4.3 Elephant-Flow Routing

Elephant flows are routed using a multi-objective ACO algorithm. Unlike mice flows, elephant flows require path delay, path risk, and path remaining bandwidth. For this reason, the algorithm constructs a set of candidate routes and evaluates them by multiple criteria before selecting a multipath routing plan.

Pheromone Initialization. Before the ACO starts, a delay-based Dijkstra algorithm is computed between the source and the destination nodes of the elephant flow. This route is used to initialize the pheromone matrix. Hence, the search does not start from a completely uniform pheromone distribution, but receives a meaningful initial direction.

This initialization improves convergence of the ACO algorithm and reduces the probability of spending many iterations on bad routes. At the same time, it does not restrict the ants to only the seeded route, since the ants can still discover another and better alternatives.

Path Construction by Ants. After initialization, ants construct candidate routes between the source and destination nodes. Let τ_{ij} be the pheromone level

of edge (i, j) and let η_{ij} be the heuristic desirability of that edge. Then the probability of moving from node i to node j is defined as

$$P_{ij} = \begin{cases} \dfrac{\tau_{ij}^{\alpha}\eta_{ij}^{\beta}}{\sum_{k\in\mathcal{N}(i)}\tau_{ik}^{\alpha}\eta_{ik}^{\beta}}, & \text{if } j \in \mathcal{N}(i), \\ 0, & \text{otherwise}, \end{cases} \tag{30}$$

where $\mathcal{N}(i)$ is the set of feasible neighbors of node i, and α and β determine the influence of pheromone and heuristic information, respectively.

In the proposed method, the heuristic desirability of an edge is defined as

$$\eta_{ij} = \frac{1}{d_{ij}} \cdot \frac{1}{r_{ij}+\varepsilon} \cdot b_{ij}^{\mathrm{rem}} \cdot \frac{1}{(1+\mathrm{dist}(j,s'))^{\psi}}. \tag{31}$$

where b_{ij}^{rem} is the remaining bandwidth of edge (i, j), d_{ij} is the delay, r_{ij} is the edge risk, $\varepsilon > 0$ is a small positive constant preventing division by zero, and $\mathrm{dist}(j, s')$ is the hop distance from node j to the destination node. The last factor reduces the probability of long detours during path construction. Therefore, the ACO prefers edges that are fast, reliable, and still have enough available capacity.

Multi-objective Path Evaluation. For every candidate route $\mathcal{P}_q$, the following objective vector is evaluated:

$$F(\mathcal{P}_q) = (f_1(\mathcal{P}_q), f_2(\mathcal{P}_q), f_3(\mathcal{P}_q)), \tag{32}$$

where

$$f_1(\mathcal{P}_q) = D(\mathcal{P}_q), \qquad f_2(\mathcal{P}_q) = R(\mathcal{P}_q), \qquad f_3(\mathcal{P}_q) = \frac{1}{B(\mathcal{P}_q)}. \tag{33}$$

Thus, the algorithm simultaneously minimizes path delay, path risk, and inverse path bottleneck remaining bandwidth.

Archive Update and Diversity Preservation. After candidate routes are generated, they are stored in the archive. Then all routes from archive are filtered by nondominated sorting. Only Pareto-optimal routes are stored in the archive. If the number of stored solutions is larger than the archive size, crowding distance is used in order to preserve diversity among the stored routes.

Therefore, the archive contains not only high-quality routes, but also routes representing different trade-offs between delay, risk, and inverse bottleneck remaining bandwidth.

Pheromone Update. Pheromone values are updated globally according to the quality of the archived routes. The update rule is

$$\tau_{ij} = \begin{cases} (1-\rho)\tau_{ij} + \dfrac{\Delta\tau_{ij}}{|\mathcal{P}_q|}, & \text{if } (i,j) \in \mathcal{P}_q, \\ (1-\rho)\tau_{ij}, & \text{otherwise}, \end{cases} \tag{34}$$

where ρ is the pheromone evaporation parameter,

$$\Delta\tau_{ij} = \frac{1}{score}, \qquad score = \frac{\widetilde{d}_{ij} + \widetilde{r}_{ij} + \widetilde{b}_{ij}^{-1}}{3} \tag{35}$$

where $\widetilde{d}_{ij}$ is the normalized delay, $\widetilde{r}_{ij}$ is the normalized risk, $\widetilde{b}_{ij}^{-1}$ is the normalized inverse remaining bandwidth.

If the current archive contains too few distinct routes, the algorithm interprets this situation as a loss of exploration. In this case, the pheromone evaporation rate is increased. A larger value of ρ weakens the influence of over-dominant routes and encourages the ants to explore alternative directions of the search space.

As well, to prevent early convergence, pheromone values are bounded:

$$\tau_{\min} \leq \tau_{ij} \leq \tau_{\max}. \tag{36}$$

The upper bound prevents a single route from dominating the entire search, whereas the lower bound prevents potentially useful edges from being completely ignored.

Final Path Subset and Proportional Allocation. After the search phase, the algorithm does not use all archived routes. Instead, it selects a subset of routes that are both high-quality and enough different from each other. For this purpose, a simple diversity filter is applied in the final selection stage in order to avoid several nearly identical routes.

Suppose that the selected routes have bottleneck remaining bandwidth values $B_1, \ldots, B_{m_q}$. Then the traffic share assigned to route r is defined as

$$y_q^{(r)} = \frac{B_r}{\sum_{j=1}^{m_q} B_j}, \qquad \sum_{r=1}^{m_q} y_q^{(r)} = 1. \tag{37}$$

Thus, the demand of elephant flow is distributed proportionally to available path capacity. This makes the final allocation simple and feasible.

4.4 Pseudocode of TFPACO

The overall logic of the proposed method for mice is summarized in Algorithm 1, and for elephant in Algorithm 2.

4.5 Main Features of the Proposed Method

The main difference between TFPACO and the baseline methods is that it combines division into classes of mice and elephants, multi-objective ACO, and multipath allocation in the SDN. In comparison with Dijkstra and multi-objective Dijkstra algorithms, the proposed method does not reduce the routing problem to a single scalar path cost. In comparison with TSACO, it uses multi-objective optimization and takes into account three criteria instead of just one. In comparison with the Motlagh algorithm, it divides traffic into classes and use more criteria in multi-objective ACO.

Algorithm 1. TFPACO: mice

Require: flow = (src, dst, demand, lifetime, class=mice), network graph G
Ensure: route plan
  Compute shortest path P using Dijkstra with weight $w_{ij} = d_{ij} + \phi r_{ij}$
  if P is feasible **then**
    return single-path routing $\{(P, demand)\}$
  else
    return error
  end if

Algorithm 2. TFPACO: elephant

Require: flow = (src, dst, demand, lifetime, class=elephant), network graph G
Ensure: route plan
  Compute shortest path P using Dijkstra with weight $w_{ij} = d_{ij}$
  Initialize pheromone matrix τ
  Update pheromone on edges of path P
  Initialize archive $A = \emptyset$
  MOACO(flow, network graph G, pheromone matrix τ, archive A)
  Select diverse subset of routes from archive A
  Assign traffic shares y_r proportionally to bottleneck bandwidth
  return multipath routing plan $\{(P_1, y_1 \cdot demand),\ \ldots,\ (P_k, y_k \cdot demand)\}$

Algorithm 3. TFPACO: MOACO

Require: flow = (src, dst, demand, lifetime, class=elephant), network graph G, pheromone matrix τ, archive A
Ensure: archive A
  for iter = 1,. . . ,T **do**
    Generate candidate paths using ants
    for each candidate path P **do**
      Compute objectives: $f_1(P) = D(P),\ f_2(P) = R(P),\ f_3(P) = \frac{1}{B(P)}$
    end for
    Merge candidates with archive A
    Perform nondominated sorting
    Keep Pareto-optimal solutions
    if archive size is too large **then**
      Apply crowding-distance selection
    end if
    if route diversity in A is too low **then**
      Increase evaporation coefficient ρ
    end if
    Update pheromone values
    Enforce pheromone bounds: $\tau_{\min} \leq \tau_{ij} \leq \tau_{\max}$
  end for

5 Simulation Model and Experimental Setup

5.1 Experimental Setup

To evaluate the performance of the proposed TFPACO algorithm, we implemented a discrete-time simulation environment in Python [9]. The network is modeled as a graph $G = (V, E)$, consisting of $|V| = 100$ nodes and a randomly generated connected topology. Each link $(i, j) \in E$ is characterized by a set of attributes: base delay $d_{ij}^{(0)}$, bandwidth b_{ij}, and base packet loss probability $p_{ij}^{(0)}$.

To reflect real-world network dynamics and the impact of congestion on QoS, we utilize non-linear cost functions for delay and loss. As noted in recent literature [6,8], link performance degrades as utilization increases. We define the link delay $d_{ij}^{(t)}$ and loss probability $p_{ij}^{(t)}$ at step t as [10]:

$$d_{ij}^{(t)} = d_{ij}^{(0)} \left(1 + k \cdot u_{ij}(t)^z\right), \qquad p_{ij}^{(t)} = p_{ij}^{(0)} + h \cdot u_{ij}(t)^\gamma, \tag{38}$$

where

$$u_{ij}(t) = \frac{b_{ij}^{\text{used}}(t)}{b_{ij}(t)} \tag{39}$$

represents the current link utilization. We set $k = 4.0$ with $z = 2.0$ for the delay model, and $h = 0.15$ with $\gamma = 3.0$ for the loss model, ensuring a sharp QoS degradation beyond the 90% utilization threshold.

The simulation is a sequence of flows

$$q = (s, s', \lambda, \xi, c), \tag{40}$$

where s and s' are randomly sampled nodes, λ is the bandwidth demand in kbps, ξ is the flow duration in discrete steps, and c is the traffic class.

Flows are categorized into:

- **Mice Flows:** Short-lived (2–8 steps), low-bandwidth (100–800 kbps), representing 80% of the total flow count.
- **Elephant Flows:** Long-lived (20–60 steps), high-bandwidth (2000–8000 kbps), representing 20% of the count but contributing approximately 90% of the total volume.

The simulation runs for $T = 50$ steps, with 50 new flows generated at each step, creating a high-pressure scenario to test the scalability and efficiency of TFPACO algorithm.

The performance of routing algorithms is measured through a comprehensive set of QoS metrics and network indicators:

- **Blocked Flow/Demand Ratio:** The percentage of flows and total bandwidth demand of them that the controller couldn't= process due to resource constraints.
- **p95 Experienced Delay (d_{95}^{exp}):** The 95th percentile of the actual latency experienced by admitted flows, focusing on tail-latency.

- **p5 Packet Delivery Ratio (PDR_5^{exp}):** The lower 5th percentile of PDR, highlighting the reliability of the most vulnerable paths.
- **Jain's Fairness Index:** A measure of load distribution across the network, where 1.0 indicates perfectly balanced utilization.
- **Congestion:** The percentage of simulation time during which at least one edge is heavily overloaded (above 90% capacity).

5.2 Experimental Results

The comparison of the TFPACO algorithm with the Dijkstra, the Multi-Objective Dijkstra, TSACO, and the Motlagh algorithms is presented in Tables 1 and 2 and Figs. 1–5.

As illustrated in Table 1, the Dijkstra algorithm has the lowest latency for admitted flows, but at the cost of very high demand blocking rate (49.86%), since it does not take into account the available bandwidth. The MO-Dijkstra variants [6] reduce the demand blocking to some extent (approx. 30%), since they take into account the available bandwidth. However, they still fail to support high demand requirements.

In contrast, TSACO, Motlagh and TFPACO significantly reduce blocking. However, the computational cost of the TSACO and Motlagh MOACO is extremely high. The proposed TFPACO algorithm achieves a 100% admission rate (0.00% blocking), while maintaining a much smaller runtime (see Table 2) and average delay (see Fig. 1) than TSACO and Motlagh, and a much smaller mice average delay (see Fig. 2) than Motlagh, demonstrating the benefits of separating traffic classes and multi-objective ACO.

Table 1. Comparison of Routing Algorithms: QoS and Reliability Metrics

Algorithm	Flow Blk., %	Dem. Blk., %	d_{95}^{exp} (ms)	d_{95}^{mice} (ms)	PDR_5^{exp}, %
Dijkstra (Delay)	15.12	49.86	28.84	25.78	82.10
MO-Dijkstra (Cfg1)	7.88	33.30	40.25	33.36	89.23
MO-Dijkstra (Cfg2)	8.28	34.08	38.75	32.22	89.55
MO-Dijkstra (Cfg3)	6.72	29.10	46.69	37.90	91.38
TSACO	1.92	0.88	160.36	28.35	80.63
Motlagh MOACO	0.00	0.00	261.40	256.50	66.74
TFPACO	**0.00**	**0.00**	**71.49**	**27.58**	**86.28**

(See Figs. 3 and 4)

Table 2. Network State and Computational Overhead Metrics

Algorithm	Runtime (ms)	Max Util., %	Congestion, %	Jain Index
Dijkstra (Delay)	2,572.2	99.90	100.0	0.181
MO-Dijkstra (Cfg1)	4,947.0	99.88	100.0	0.211
MO-Dijkstra (Cfg2)	5,020.1	99.90	100.0	0.212
MO-Dijkstra (Cfg3)	5,086.7	99.85	100.0	0.213
TSACO	1,109,019.0	99.83	92.0	0.493
Motlagh MOACO	907,056.6	99.90	100.0	0.344
TFPACO	**103,602.4**	**99.90**	**70.0**	**0.438**

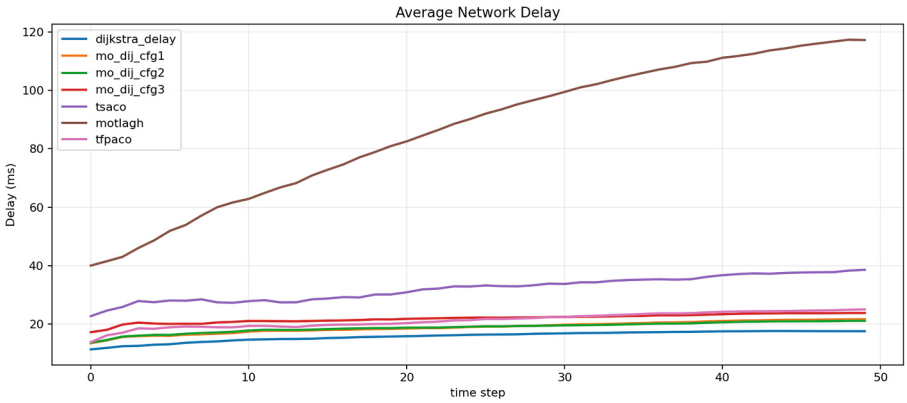


Fig. 1. Average delay over simulation time for the compared algorithms.

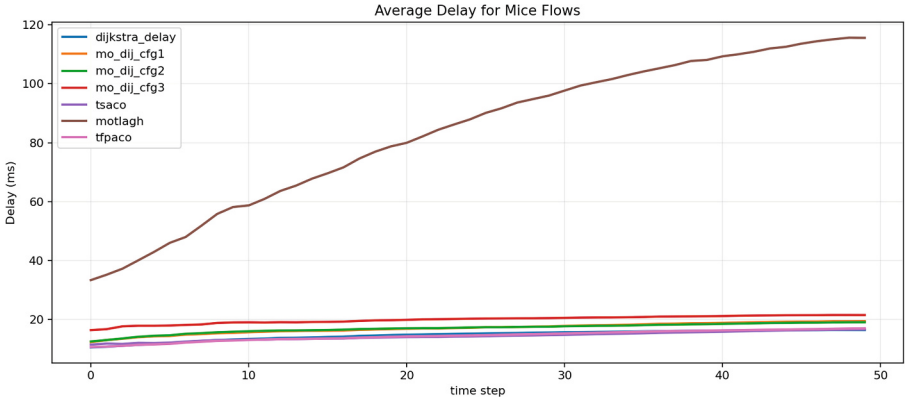


Fig. 2. Average mice delay over simulation time for the compared algorithms.

Fig. 3. Packet delivery ratio over simulation time for the compared algorithms.

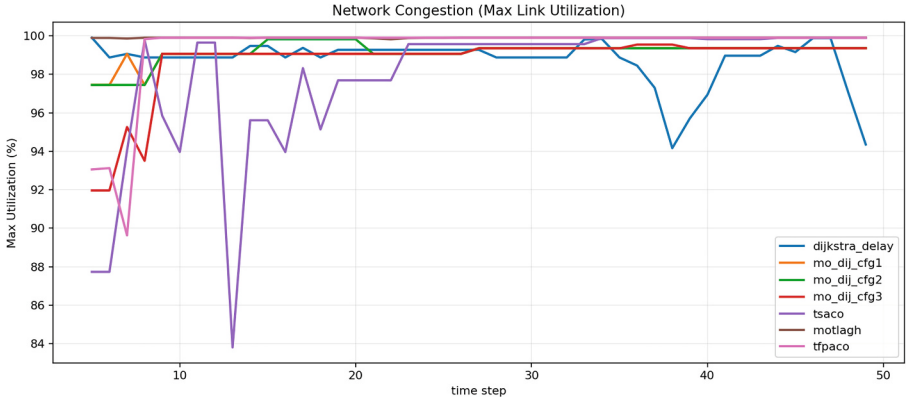


Fig. 4. Maximum link utilization over simulation time for the compared algorithms.

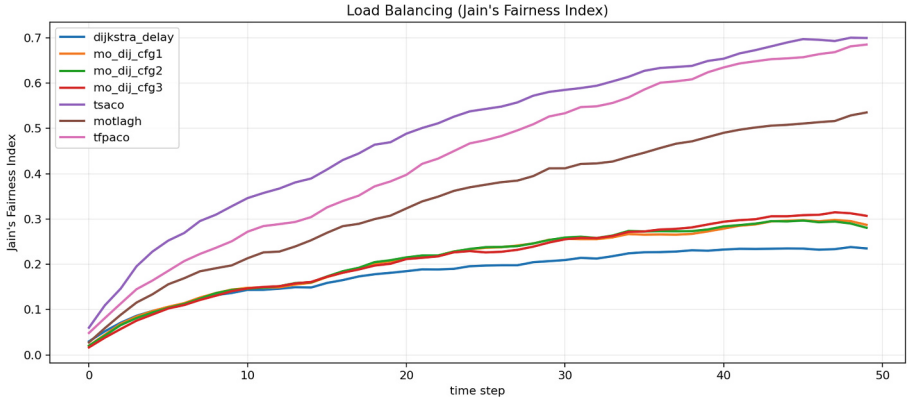


Fig. 5. Network load dynamics over simulation time for the compared algorithms.

6 Conclusion

The paper addressed the problem of traffic management in software-defined networks with QoS criteria and dividing traffic flows in two classes. Since mice flows and elephant flows have different requirements, it is not reasonable to use a common routing strategy. To deal with this problem, we proposed the TFPACO algorithm, which combines the Dijkstra algorithm for mice flows with a multi-objective ant colony optimization algorithm for elephant flows. For elephant flows, the proposed method takes into account path delay, path risk, and inverse bottleneck remaining bandwidth, where the archive of nondominated candidate paths is used for further multipath allocation of the traffic flows.

The proposed algorithm was evaluated in a discrete-time simulation environment with dynamic edge parameters and was compared with the Dijkstra algorithm, multi-objective Dijkstra algorithm, TSACO, and the Motlagh algorithm. The results show that TFPACO is effective in selecting routes for both traffic types and in improving overall routing performance. In particular, the algorithm provides lower latency and better traffic distribution, while avoiding flow and demand blocking under the considered simulation parameters. These results indicate that the proposed method is a promising solution for SDN traffic engineering in networks with different QoS requirements.

References

1. Xia, W., Wen, Y., Foh, C.H., Niyato, D., Xie, H.: A survey on software-defined networking. IEEE Commun. Surv. Tutorials **17**(1), 27–51 (2015)
2. Xu, C., Qin, D., Song, F.: A survey of SDN traffic management research. In: 11th International Conference on Communications, Circuits and Systems (ICCCAS), pp. 231–236 (2022)
3. Tache, M.D., Păscuțoiu, O., Borcoci, E.: Optimization algorithms in SDN: routing, load balancing, and delay optimization. Appl. Sci. **14**(14), 5967 (2024)
4. Wang, C., Zhang, G., Xu, H., Chen, H.: An ACO-based link load-balancing algorithm in SDN. In: 7th International Conference on Cloud Computing and Big Data (CCBD), pp. 214–218 (2016)
5. Wang, C., Zhang, G., Chen, H., Xu, H.: An ACO-based elephant and mice flow scheduling system in SDN. In: 2nd IEEE International Conference on Big Data Analysis (ICBDA), pp. 859–863 (2017)
6. Choudhary, A., Kang, S.S., Singla, S.: Multi-objective Dijkstra algorithm for enhancing QoS in SDN through balanced routing. In: 7th International Conference on Circuit Power and Computing Technologies (ICCPCT), pp. 1169–1175 (2024)
7. Taghavi Motlagh, S., Ibrahim, A., Shah Heydari, S., El-Khatib, K.: Multi-objective optimization for managing disruption risk in SDN. In: 20th International Conference on the Design of Reliable Communication Networks (DRCN), pp. 1–8 (2024)
8. Shreyas, J., Jumnal, A., Udayaprasad, P.K., Rekha, C., Askar, S.S., Abouhawwash, M.: Bio-inspired ACO-based traffic aware QoS routing in software defined Internet of Things. Appl. Artif. Intell. **38**(1), 2371739, 1–26 (2024)

9. Safonenko, A.: TFPACO algorithm LNCS. https://github.com/NillFin/TFPACO-algorithm-LNCS-26. Accessed 30 Mar 2026
10. Qiu L., Yang Y.R., Zhang Y., Shen S.: On selfish routing in internet-like environments. In: Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (ACM SIGCOMM), pp. 151–162 (2003)

Multithreaded Pipeline Parallelism for Distributed LLM Inference Without Source Code Modification

Victor Koloskov(✉) and Anton Pershin

St Petersburg State University, St Petersburg 199034, Russia
st107068@student.spbu.ru, a.pershin@spbu.ru

Abstract. Modern large language models (LLMs) reach tens or even hundreds of billions of parameters, making it impossible to fit them into the memory of a single GPU. To perform inference for such models, multiple accelerators are used, across which the model is distributed. However, many current distributed inference frameworks require modifications to the model's source code or adaptation of its architecture. Tools that do not require model changes are typically oriented towards using tensor parallelism within a single node. In the case of inter-node interaction, inference performance with such partitioning can significantly decrease due to frequent data exchange between accelerators and limited network bandwidth. One way to reduce inter-node communication is the use of pipeline parallelism. However, existing solutions rarely provide the ability to use pipeline parallelism for distributed inference without modifying the model code. This paper presents a concept for partitioning a model between accelerators by modifying its layers without affecting the forward pass function itself. A realization of pipeline parallelism with computational overlap between graphics accelerators is also presented. The proposed approaches allow for improved distributed inference performance without changing the LLM source code.

Keywords: Large Language Models (LLMs) · Distributed Inference · Pipeline Parallelism

1 Introduction

Currently, large language models (LLM) may require hundreds of gigabytes of accelerator memory for deployment and inference. For example, the Llama 3.1-405B model [1] is the flagship model of the Llama family and requires approximately 810 GB in FP16 format just for weight storage. The DeepSeek-V3 model [2] contains 671 billion parameters and requires approximately 1350 GB of memory. At the same time, the NVIDIA H200 [3], currently the accelerator with the largest memory capacity, is equipped with 141 GB HBM, which is insufficient to run Llama 3.1-405B or DeepSeek-V3 on a single accelerator. To solve this problem, the model is split between several interconnected accelerators, adding the

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 464–480, 2027.
https://doi.org/10.1007/978-3-032-30530-5_29

necessary data exchange operations between GPUs. There are two main ways of model partitioning: intra-layer and inter-layer.

In intra-layer partitioning, each layer is divided among several GPUs. This allows for the parallel execution of a single layer's computations, which is especially effective for layers with a large number of parameters. The main advantage of this approach is the ability to evenly distribute the computational load between accelerators. However, this method requires frequent exchange of intermediate activations between GPUs, making its efficiency highly dependent on network bandwidth. Therefore, this partitioning method is typically most effective within a single computing node. In LLM distributed computing tasks, this approach is usually implemented as **tensor parallelism**.

In inter-layer partitioning, each GPU receives its own group of layers for which it performs computations. Data is passed sequentially between accelerators as it moves through the model's layers. The main advantage of this approach is lower dependence on network bandwidth, allowing it to be used for inter-node interaction. However, the computational load between accelerators may be distributed unevenly and requires correct execution queue management to reduce GPU idle time. In LLM distributed computing tasks, this approach is implemented as **pipeline parallelism**.

Existing frameworks for distributed inference [4–10] typically support both intra-layer and inter-layer model partitioning. However, leveraging these tools typically necessitates modifying the model's source code or adapting its architecture[1], which requires deep knowledge of the model's internal structure. Solutions that do not require such modifications usually rely on tensor parallelism within a single node, a strategy that often proves inefficient for inter-node scaling due to high communication overheads.

Pipeline parallelism is significantly less common as a ready-made solution for distributed inference, and existing implementations do not always provide high efficiency. This is because, in naive implementations, an accelerator cannot start processing a new batch until the previous one is fully processed at all stages of the pipeline, resulting in GPUs being idle for a significant amount of time.

In this work, we present the concept of **placing a model on several accelerators by substituting model layers** which preserves the original forward pass function and eliminates the need for source code modification (Sect. 3). We also present an **implementation of pipeline parallelism for distributed inference with input batch splitting into several microbatches**, which reduces GPU idle time (Sect. 4).

2 Related Work

Pipeline parallelism was first examined in detail in the work by Y. Huang et al. [12], where the GPipe library for neural network training was introduced. The

[1] Source code adaptation is defined as the modification of the forward pass implementation and layer computation logic necessary to distribute the model across multiple accelerators for inference.

key innovation was splitting the batch into several microbatches. This made it possible to reduce the "bubble"—the idle period of computing resources occurring when there is no load on the GPU—and thus increase the speed of the forward and backward passes. The next improvement was proposed in the work by A. Harlap et al. [13] with the Pipedream framework. The paper presented a new pipeline schedule that further reduced GPU idle time during model training. Subsequent works on pipeline parallelism [14–17] primarily aimed to improve scheduling to minimize accelerator idle time. Today, pipeline parallelism is implemented in several popular frameworks. Below is their description and comparative analysis.

2.1 Megatron-LM [4]

The Megatron-LM framework was first presented in the work by M. Shoeybi et al. [4] as a solution for LLM training using tensor and data parallelism. Today, it also supports pipeline parallelism [18] and has tight integration with NVIDIA accelerators. However, the framework is more often used for model training and requires adapting the model's source code to its architecture.

2.2 MindSpore [5]

MindSpore is an open and flexible computing platform for developing and executing AI models, created by Huawei. The platform is distributed with fully open source code and, besides GPUs, can run on Ascend NPU processors. MindSpore supports several types of parallelism and can automatically select optimal model partitioning strategies between accelerators. However, using the platform requires model source code adaptation.

2.3 Colossal-AI [6]

Initially, Colossal-AI was developed as a solution for LLM training on multiple GPUs with support for various types of parallelism: data, pipeline, tensor, sequence parallelism, and for working with heterogeneous systems. Subsequently, Colossal-AI also began to be used for distributed inference. For standard models from the Transformers library, additional adaptation is generally not required. In the absence of support for a specific model, it is necessary.

2.4 vLLM [7]

The vLLM framework is designed for high-performance LLM inference. The system's primary goal is to increase throughput and reduce latency when serving a large number of requests. To achieve this, optimizations for memory management, batching, and efficient KV-cache management during generation have been implemented.

Furthermore, vLLM is tightly integrated with the HuggingFace Transformers library and provides broad coverage of models used in distributed inference. If a

model is not supported, it can still be used through the "Transformers modeling backend". However, the model must satisfy certain requirements [11].

2.5 TensorRT-LLM [8]

TensorRT-LLM is a system for high-performance LLM inference on NVIDIA GPUs. It is based on the NVIDIA TensorRT library and utilizes low-level optimizations such as kernel fusion, graph optimization, and specialized CUDA kernels to accelerate inference. TensorRT-LLM supports several types of parallelism but requires converting the model into a format compatible with TensorRT.

2.6 DeepSpeed [9]

The DeepSpeed system provides the DeepSpeed-Inference tool [19], designed to accelerate the inference of large language models. To improve text generation efficiency, the library also implements a specialized engine, DeepSpeed-FastGen [20], designed to accelerate autoregressive token generation. DeepSpeed-Inference fully supports tensor parallelism for any model without additional code adaptation by replacing model layers with parallel ones. It also supports ZeRO-Inference technology [21], which allows for the asynchronous offloading of model layers to RAM or disk and loading them back as needed. For pipeline parallelism, model adaptation is still required.

2.7 Accelerate [10]

Accelerate was developed by Hugging Face for distributed LLM training and inference without the need to adapt the model. It can automatically distribute the model between accelerators and work with it as if it were not partitioned. The main disadvantage is that, in a strict sense, the framework does not implement a specific type of parallelism but only shards the model, which can result in lower inference throughput.

2.8 Summary Comparison

Table 1 presents a summary comparison of existing solutions for distributed LLM inference. The analysis shows that currently, there is no tool that simultaneously provides efficient use of pipeline parallelism and does not require significant model adaptation for a specific framework.

3 Proposed Method for Model Partitioning

The main problem with inter-layer model partitioning is the need to modify the model's forward pass function to ensure its operation on multiple accelerators. This work proposes adapting the approach used in the DeepSpeed-Inference system [19] with tensor parallelism to implement pipeline parallelism. The core

Table 1. Comparison of existing distributed inference frameworks

Framework	Key Characteristics	Source Code Adaptation
Megatron-LM	tight integration with NVIDIA GPUs	requires model source code adaptation
MindSpore	automatically selects optimal model partitioning; supports Ascend NPU	requires model source code adaptation
Colossal-AI	support for heterogeneous systems	supports standard Transformers model; requires adaptation if a specific model is not supported
vLLM	optimizations for high-performance inference; tight integration with Transformers library	Supports a wide range of models; unsupported Transformers models can be used via the Transformers backend, subject to certain requirements
TensorRT-LLM	uses low-level optimizations for NVIDIA GPU acceleration	requires model conversion to TensorRT-compatible format
DeepSpeed	multiple inference optimizations; ZeRO-Inference	tensor parallelism is automatic; pipeline parallelism requires adaptation
Accelerate	Ease of use; tight integration with Transformers library	requires no code adaptation, though inference throughput is generally lower

idea is to replace standard model layers with specialized parallel layers, which allows avoiding rewriting the forward pass function and limiting changes only to layer replacement. An additional advantage of this approach is that, in pipeline parallelism, each layer is executed on only one device, so its computational logic does not require modification. However, this property simultaneously creates an additional challenge. Even when partitioning the model between accelerators, it is necessary to execute the forward pass function on each GPU and pass intermediate data between pipeline stages.

To solve this task, three types of layers are introduced: **Compute**, **Fake**, and **Boundary**. The **Compute** layer performs the main computations of the corresponding model layer on the GPU. The **Boundary** layer is responsible for executing computations in the final layer of the current stage and transmitting the output to the subsequent accelerator, as well as receiving input data from the preceding stage. The **Fake** layer contains no trainable parameters and is used to generate dummy zero-value output data necessary for executing the forward pass function on a GPU where the current stage's computations are not being performed. This design ensures that the forward pass function requires no modification, as the substitution operates at the module level rather than altering the underlying computation logic.

Figure 1 shows an example of model partitioning across multiple GPUs. In this case, the model is distributed among several accelerators, yet for the forward pass function, it remains logically complete on each device. Final **Boundary** layer is additionally required to propagate up-to-date output data to all accelerators. This is especially important in iterative inference procedures, such as autoregressive generation, where the first pipeline stage must receive the actual output produced in the previous step.

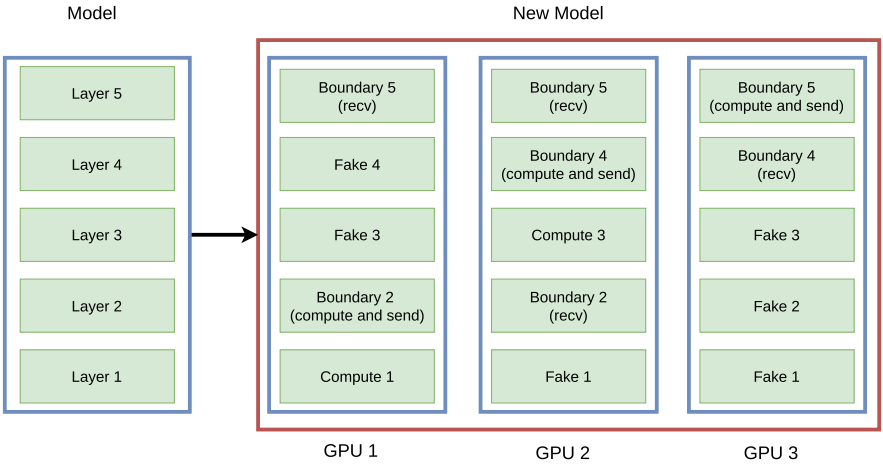


Fig. 1. Example of partitioning a five-layer model between several GPUs.

Figure 2 illustrates the forward pass process of a partitioned model. The input batch arrives at all accelerators simultaneously. On GPU 2 and GPU 3, based on the input data, necessary intermediate values are formed in Fake-type layers up to the Boundary layer, where execution is paused until output data are received from the previous stage. At the same time, standard computations are performed on Compute layers on GPU 1. At the Boundary layer, activations of the last layer of the current stage are calculated, and then the data is passed to GPU 2.

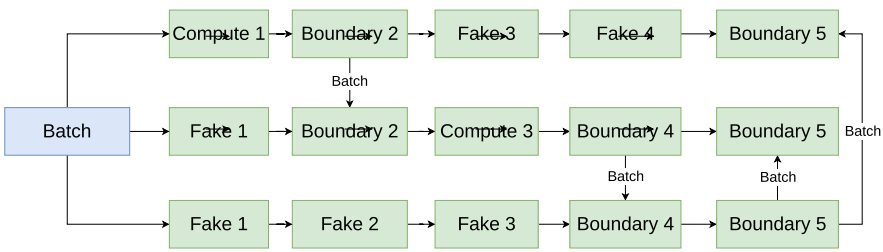


Fig. 2. Example of the forward pass of a partitioned model.

After receiving input data on GPU 2, the next stage of the pipeline is executed, and upon its completion, the batch is passed to GPU 3. On GPU 3, the remaining part of the forward pass is executed, after which the final layer's output data is passed to all accelerators.

To simplify the layer replacement process in the model, the **Generator** mechanism is proposed. Initially, the model in each process is placed in RAM. Each

model layer is passed to a corresponding Generator, which, depending on the chosen strategy, replaces the original layer with one of the types: Compute, Fake, or Boundary, after which the modified layer is placed on the GPU. Each process uses its own sequence of generators defining the rules for model layer transformation.

A natural concern with the layer substitution approach is the handling of non-linear architectural patterns such as skip connections and Mixture-of-Experts layers. As a general design principle, the partitioning strategy should aim to avoid splitting such non-linear constructs across pipeline stages: the granularity of partitioning is to be chosen such that these architectural features remain encapsulated within a single indivisible block. This principle is adopted in the present work and ensures that no skip connection crosses a stage boundary.

In cases where such splitting is unavoidable, the framework can be extended to support cross-stage data flow through the introduction of additional Fake-layer variants. Under this scheme, the auxiliary tensors required by the consuming stage are transmitted from the producing stage prior to the start of the next pipeline stage and retained at the receiving stage until the point in the forward pass at which they are consumed. A detailed design and evaluation of this extension is left for future work.

Thus, adapting a new model is reduced to implementing a corresponding Generator that defines the layer replacement policy and establishes a uniform rule across all processes for generating Fake layer outputs based on the input batch. In essence, this reduces model support to a declarative description of the model's structure, without altering its underlying computational logic. If necessary, the architecture can be extended by adding new data transfer strategies between GPUs or additional Fake-category layer types.

4 Multithreaded Pipeline Parallelism

The proposed implementation of pipeline parallelism has a significant drawback. GPUs that have finished their stage's computations cannot start processing the next batch because they must wait for output data at the final layer (Fig. 2). As a result, accelerators may stand idle, leading to reduced inference throughput.

To solve this problem, we propose using multithreading within the pipeline. Similarly to GPipe [12], we split the original batch into several microbatches. Each microbatch is processed by a separate execution thread. After completing the current stage's computations, the thread passes the calculated activations to the next pipeline stage and suspends execution. Control is then passed to the next thread, which begins processing another microbatch (Fig. 3). This ensures overlapping of computations across different pipeline stages and reduces GPU idle time.

Figure 4 shows a comparison of pipeline parallelism schedules without and with multithreading.

The multithreading functionality is implemented using Python's standard threading library. A natural concern is Python's Global Interpreter Lock (GIL),

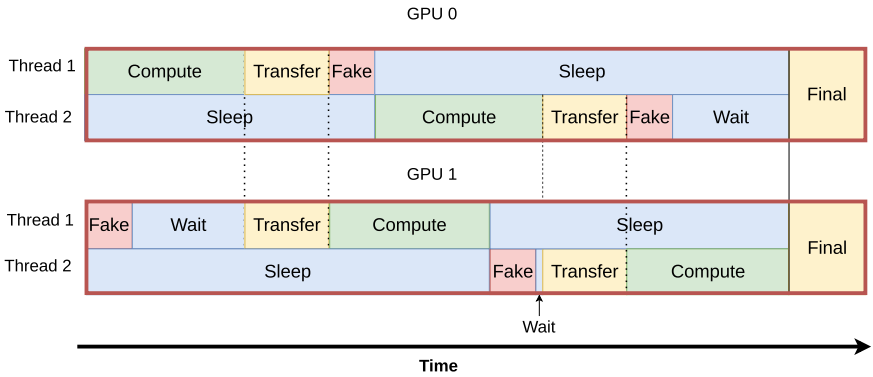


Fig. 3. Example of multithreaded pipeline operation with two GPUs and two threads.

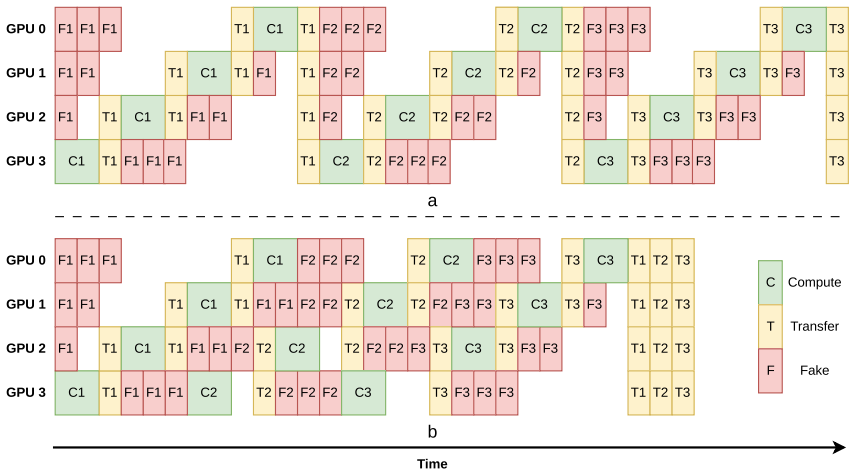


Fig. 4. Comparison of pipeline schedules without multithreading (a) and with multithreading (b).

which prevents multiple threads from executing simultaneously within a single process. However, this limitation does not affect performance in our case: each process is responsible for a separate pipeline stage and contains one thread per microbatch. At any given moment, only one thread is active—the others wait for the current stage to complete (Fig. 2). Thus, the scheduler architecture naturally aligns with the constraints imposed by the GIL.

5 Evaluation of Multithreaded Pipeline Parallelism

This section demonstrates the results of comparing throughput and peak GPU memory consumption for standard and multithreaded pipeline parallelism. For

the experiments, 2 computing systems were used. Table 2 presents the system characteristics, models used, and the data type in which the experiments were conducted.

It should be noted that pipeline parallelism is typically applied in systems with limited interconnect bandwidth, such as PCIe-based intra-node configurations or inter-node connections, since it requires data transfers only at the boundaries between pipeline stages. In contrast, on systems with high-bandwidth interconnects such as NVLink, tensor parallelism is generally more efficient. For this reason, hardware configurations without NVLink were selected for the experiments.

Table 2. Conducted experiments

System	Models	Data Type
2 × NVIDIA P104-100 (8 GB)	OPT-1.3B, OPT-2.7B, DebertaV2-xlarge, DinoV2-Giant	FP32
4 × NVIDIA Tesla T4 (16 GB)	OPT-6.7B, OPT-13B, BLOOM-7B1, internVIT-6B	FP16

To conduct the experiments, a customized parameter grid was established for each model, with boundaries determined by the available accelerator memory. While this constrained the maximum batch size, input sequence length, and number of output tokens, such an approach enables a precise investigation into how pipeline parallelism efficiency depends on key inference parameters.

5.1 Experiments with 2 × NVIDIA P104-100 (8 GB)

Decoder-Only Models. Figure 5 presents the results for the OPT-1.3B model. The graphs on the left show the ratio of multithreaded pipeline throughput to standard pipeline parallelism throughput under various batch sizes, input sequence lengths, and numbers of generated tokens. The graphs on the right demonstrate the optimal number of microbatches the input batch should be divided into to achieve maximum throughput. The center graphs show the ratio of peak GPU memory consumption when using multithreaded versus standard pipelines. On average, the greatest acceleration is achieved with a microbatch size of 24 and ranges from 1.10× to 1.24×. It can also be noted that as the input sequence length increases, the acceleration grows, whereas the peak memory consumption ratio decreases.

Figure 6 presents the results for the prefill stage of the OPT-1.3B model. It can be seen that the average acceleration is higher in this case, while the peak memory consumption ratio remains virtually unchanged. Based on the results shown in Figs. 5 and 6, it can be concluded that the multithreaded pipeline is most effective during the prefill stage. Furthermore, as the batch size, sequence

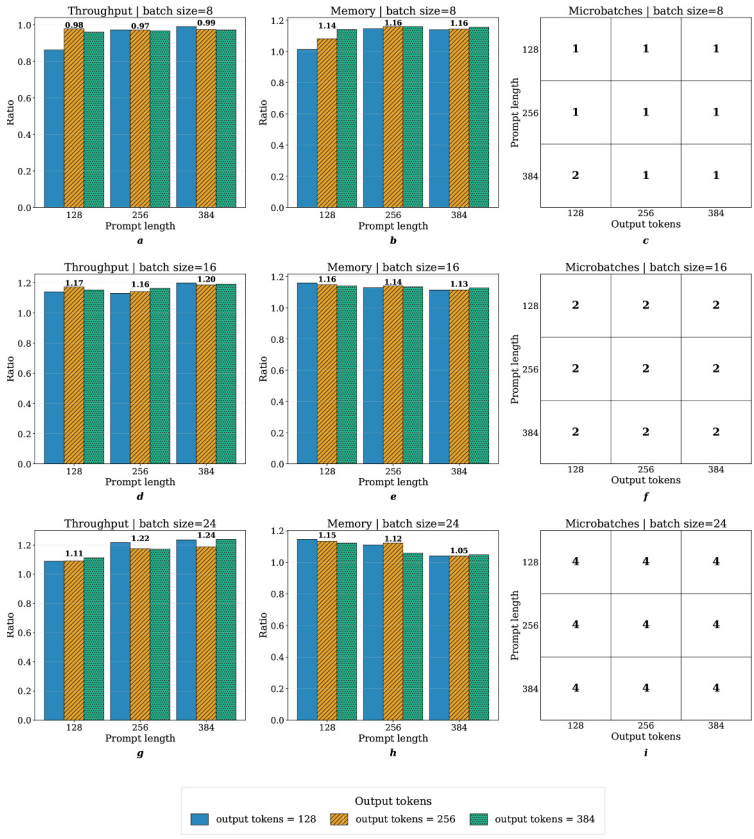


Fig. 5. Results for the OPT-1.3B model. Throughput ratio of multithreaded vs. normal pipeline (a, d, g). Peak memory consumption ratio of multithreaded vs. normal pipeline (b, e, h). Optimal number of microbatches for multithreaded pipeline (c, f, i).

length, and number of generated tokens increase, relative peak memory consumption decreases. This is explained by the fact that multithreaded mode requires some additional memory, the impact of which becomes less significant as input data parameters grow.

Figure 7 shows results for the OPT-2.7B model. It is noticeable that as the model size increases, the acceleration grows overall, reaching from 1.22× to 1.39× at a batch size of 24 at the optimal microbatch size. Simultaneously, the optimal number of microbatches increases, while the relative peak memory consumption decreases. This suggests that using a multithreaded pipeline becomes more effective as model size increases. This is due to the fact that in larger models, the computation time of individual layers is longer, increasing the opportunity for overlapping computations between pipeline stages.

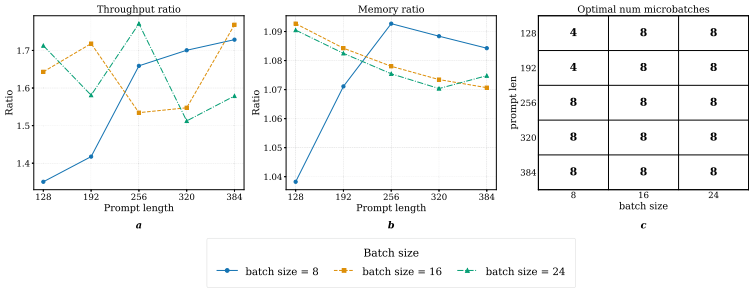


Fig. 6. Prefill stage results for the OPT-1.3B model. Throughput ratio of multithreaded vs. normal pipeline (a). Peak memory consumption ratio of multithreaded vs. normal pipeline (b). Optimal number of microbatches for multithreaded pipeline (c).

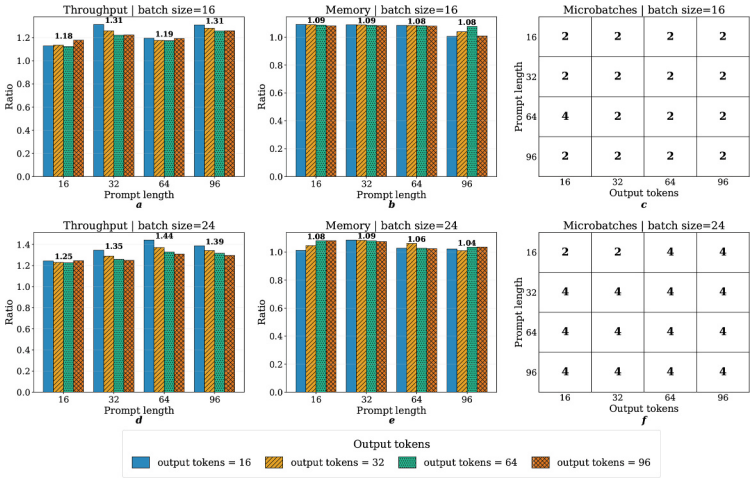


Fig. 7. Results for the OPT-2.7B model. Throughput ratio of multithreaded vs. normal pipeline (a, d). Peak memory consumption ratio of multithreaded vs. normal pipeline (b, e). Optimal number of microbatches for multithreaded pipeline (c, f).

Encoder Models. Figure 8 presents experimental results for the DebertaV2-xlarge text encoder model. The graphs show that using a multithreaded pipeline allows for an average throughput increase of up to 1.46× compared to standard pipeline parallelism. Acceleration increases with larger input sequence lengths and batch sizes. It is also observed that as these parameters increase, the peak memory consumption ratio decreases and, in some cases, can drop by nearly half. This is because the input batch is divided into smaller microbatches, reducing the memory volume required to store intermediate activations.

Figure 9 shows experimental results for the DinoV2-Giant model. The graphs show that as input batch size increases, the peak memory consumption ratio decreases. Simultaneously, with an increase in the number of microbatches, the

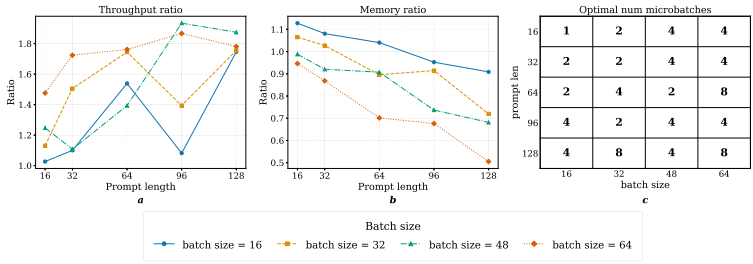


Fig. 8. Results for the DebertaV2-xlarge model. Throughput ratio of multithreaded vs. standard pipeline (a). Peak memory consumption ratio (b). Optimal number of microbatches (c).

throughput ratio of the multithreaded pipeline to the standard one increases. Using 8 microbatches results in acceleration between 1.45× and 2×. Peak memory consumption ratio also decreases, reaching as low as 0.64×.

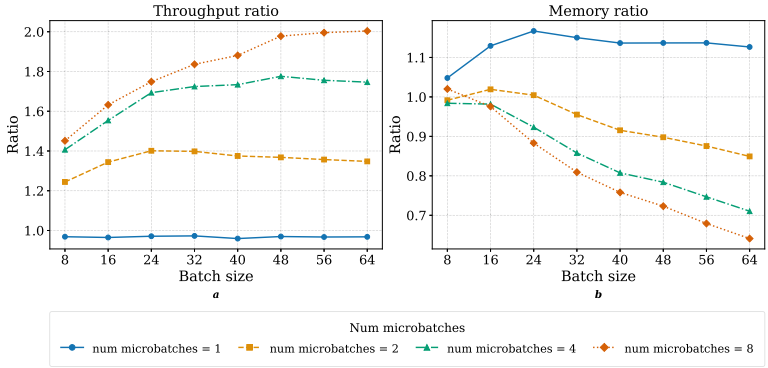


Fig. 9. Results for the DinoV2-Giant model. Throughput ratio of multithreaded vs. standard pipeline (a). Peak memory consumption ratio (b).

The results obtained for encoder architectures can be explained by the fact that their layer activation computation time is generally longer. This increases the possibility of overlapping computations between pipeline stages, leading to more efficient utilization of computing resources.

5.2 Experiments with 4×NVIDIA Tesla T4 (16 GB)

Decoder-Only Models. Figure 10 presents experimental results for the OPT-6.7B model. Acceleration at a batch size of 48 is $1.07 - 1.46\times$, while the relative peak memory consumption increases only by $1.04 - 1.08\times$.

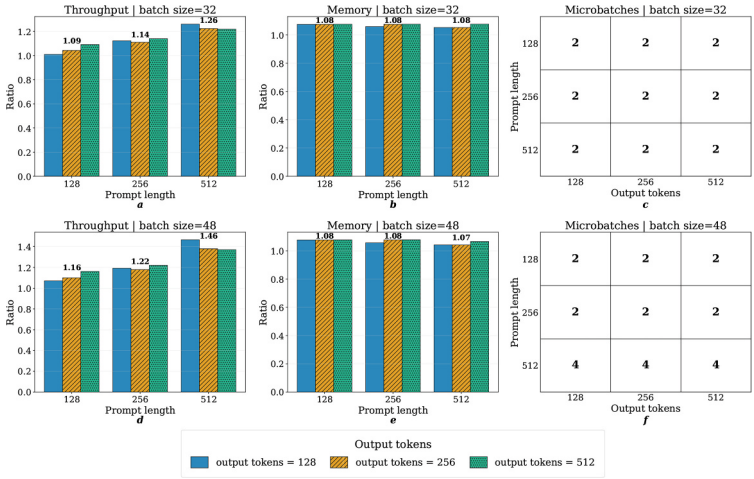


Fig. 10. Results for the OPT-6.7B model. Throughput ratio of multithreaded vs. standard pipeline (a, d). Peak memory consumption ratio (b, e). Optimal number of microbatches for multithreaded pipeline (c, f).

Figure 11 presents results for the BLOOM-7B1 model. Compared to OPT-6.7B with comparable batch parameters, sequence length, and token generation count, the average acceleration for BLOOM-7B1 is 3–4% higher. This is explained by architectural features of the model leading to heavier layer computations and a larger number of parameters. It should be noted that if the microbatch size is too small, throughput may decrease due to insufficient overlapping of computations and data transfer to the next pipeline stage.

Figure 12 presents results for the OPT-13B model. Average acceleration compared to OPT-6.7B increased by 5.7%, and the peak memory consumption ratio of the multithreaded pipeline to the standard one decreased by 2.6%. These data confirm that multithreaded pipeline efficiency grows with model size due to longer individual layer computation times and better potential for computational overlap.

A difference from the 2×P104-100 system results should be noted: on the 4×T4 system, relative peak memory consumption increases with the number of generated tokens. This is due to more complex computation distribution across multiple GPUs, resulting in increased load on individual devices, requiring more memory for multithreaded execution.

Encoder Model. Figure 13 presents results for the internVIT-6B model. Increasing the number of microbatches leads to higher acceleration, reaching $2.58 - 2.63\times$ with 8 microbatches. Simultaneously, the peak memory consumption ratio drops to $0.34 - 0.55\times$, depending on the batch size.

Comparison of results for DebertaV2-xlarge (Fig. 8), DinoV2-Giant (Fig. 9), and internVIT-6B (Fig. 13) shows that the multithreaded pipeline is especially

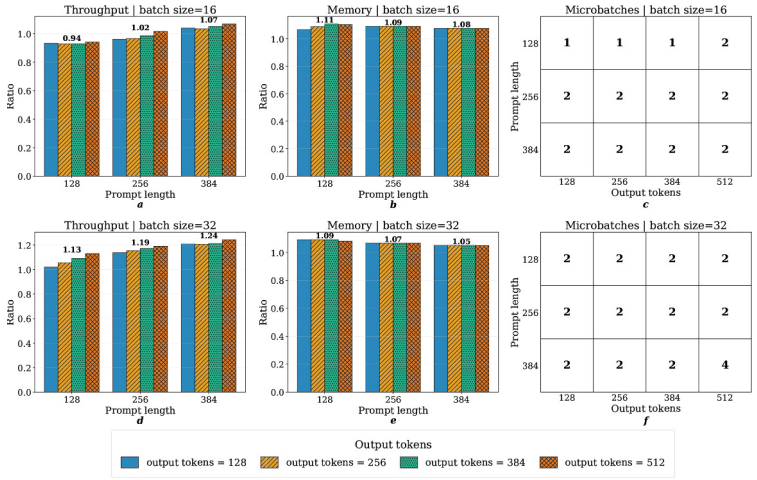


Fig. 11. Results for the BLOOM-7B1 model. Throughput ratio of multithreaded vs. standard pipeline (a, d). Peak memory consumption ratio (b, e). Optimal number of microbatches (c, f).

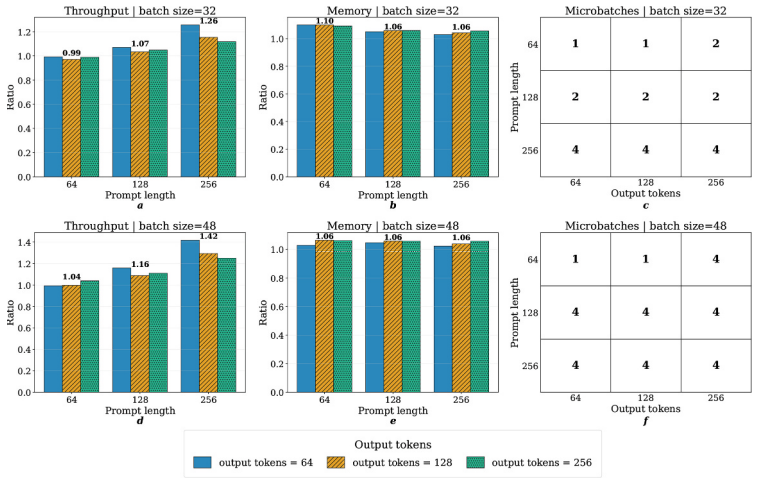


Fig. 12. Results for the OPT-13B model. Throughput ratio of multithreaded vs. standard pipeline (a, d). Peak memory consumption ratio (b, e). Optimal number of microbatches (c, f).

effective for encoder architectures. This is linked to the longer computation time for layer activations, allowing for more effective overlapping of computations across pipeline stages, thereby increasing throughput and reducing peak memory consumption.

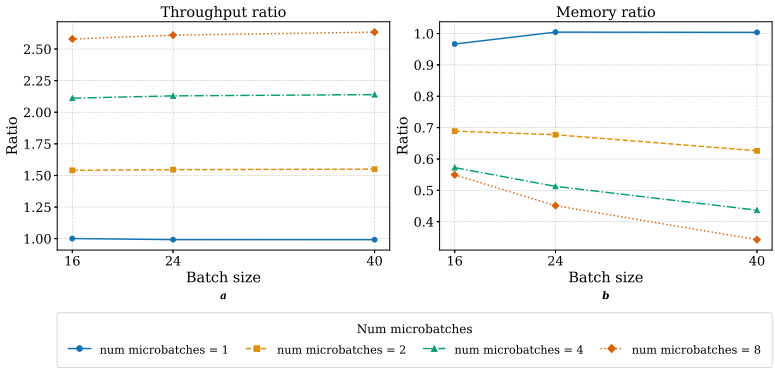


Fig. 13. Results for the internVIT-6B model. Throughput ratio of multithreaded vs. standard pipeline (a). Peak memory consumption ratio (b).

6 Conclusions and Future Work

In recent years, many powerful tools for distributed inference have been developed. However, as shown in Sect. 2, popular frameworks lack a universal solution that allows for easy adaptation of arbitrary models while ensuring high efficiency in inter-layer model partitioning. In Sect. 3, a model partitioning method based on layer substitution using generators was proposed. This approach avoids changes to the model source code and simplifies its adaptation for multi-GPU inference using pipeline parallelism. Furthermore, Sect. 4 proposed a strategy to accelerate pipeline parallelism using multithreading. Experimental results (Sect. 5) show that for Decoder-Only architectures, this approach increases inference throughput by approximately 15–30%, with an increase in peak memory consumption of about 5–15% given an optimal microbatch count. For Encoder architectures, throughput increases by 40–160% under similar conditions, while peak memory consumption decreased by up to 60%.

Limitations and Future Work. Several aspects of this study suggest directions for further investigation. First, while the use of interconnects with limited bandwidth represents a realistic deployment scenario for pipeline parallelism, evaluation on NVLink-enabled systems would help characterize the framework's behavior under high-bandwidth regimes. Second, due to GPU memory constraints of the test systems, batch sizes were limited to 48; in experiments with larger batches, memory management may begin to dominate over computational overlap, and such regimes constitute an important direction for future work. Third, the "normal" pipeline used as a baseline against the "multithreaded" pipeline corresponds to a naive sequential pipeline, which is functionally equivalent to the model sharding employed by Hugging Face Accelerate. A direct comparison with highly optimized inference systems such as vLLM, which integrate pipeline parallelism with continuous batching, paged KV-cache management, and custom CUDA kernels, would provide additional context, although

such comparisons would conflate the contribution of the parallelism mechanism itself with broader system-level optimizations and would require disentangling these factors. We plan to address all three directions in future work, alongside studying the interaction of multithreaded pipeline parallelism with other types of parallelism and optimizing pipeline scheduling to further increase throughput and reduce peak memory consumption.

References

1. Meta AI: Introducing Meta LLaMA 3.1. https://ai.meta.com/blog/meta-llama-3-1. Accessed 01 Mar 2026
2. DeepSeek-AI, Liu, A., Feng, B., Xue, B., et al.: DeepSeek-V3 Technical report. arXiv:2412.19437 (2024)
3. NVIDIA H200. https://www.nvidia.com/en-us/data-center/h200/. Accessed 25 Feb 2026
4. Shoeybi, M., Patwary, M., Puri, R., LeGresley, P., Casper, J., Catanzaro, B.: Megatron-LM: Training Multi-billion Parameter Language Models Using Model Parallelism. arXiv:1909.08053 (2019)
5. MindSpore Documentation. https://www.mindspore.cn/docs/en/r2.3.1/index.html. Accessed 21 Feb 2026
6. Li, S., et al.: Colossal-AI: a unified deep learning system for large-scale parallel training. arXiv:2110.14883 (2021)
7. vLLM GitHub. https://github.com/vllm-project/vllm. Accessed 04 Mar 2026
8. TensorRT-LLM Documentation. https://docs.nvidia.com/tensorrt-llm/index.html. Accessed 04 Mar 2026
9. DeepSpeed Official Site. https://www.deepspeed.ai. Accessed 04 Mar 2026
10. Hugging Face Accelerate. https://huggingface.co/docs/accelerate/index. Accessed 02 Mar 2026
11. vLLM Supported Models. https://docs.vllm.ai/en/latest/models/supported_models. Accessed 19 Mar 2026
12. Huang, Y., et al.: GPipe: Efficient Training of Giant Neural Networks using Pipeline Parallelism. arXiv:1811.06965 (2018)
13. Harlap, A., et a.: PipeDream: Fast and Efficient Pipeline Parallel DNN Training. arXiv:1806.03377 (2018)
14. Li, S., Hoefler, T.: Chimera: efficiently training large-scale neural networks with bidirectional pipelines. In: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, pp. 1–14 (2021)
15. Li, Z., et al.: TeraPipe: token-level pipeline parallelism for training large-scale language models. In: International Conference on Machine Learning, pp. 6543–6552. PMLR (2021)
16. Qi, P., Wan, X., Huang, G., Lin, M.: Zero Bubble Pipeline Parallelism. arXiv:2401.10241 (2023)
17. Sun, A., et al.: Seq1F1B: Efficient Sequence-Level Pipeline Parallelism for Large Language Model Training. arXiv:2406.03488 (2024)
18. Narayanan, D., et al.: Efficient Large-Scale Language Model Training on GPU Clusters Using Megatron-LM. arXiv:2104.04473 (2021)
19. Aminabadi, R.Y., et al.: DeepSpeed Inference: Enabling Efficient Inference of Transformer Models at Unprecedented Scale. arXiv:2207.00032 (2022)

20. Holmes, C., et al.: DeepSpeed-FastGen: High-Throughput Text Generation for LLMs via MII and DeepSpeed-Inference. arXiv:2401.08671 (2024)
21. ZeRO-Inference: Democratizing Massive Model Inference. https://www.deepspeed.ai/2022/09/09/zero-inference.html. Accessed 05 Mar 2026

Multilayer Computing Systems

Evgeniy Ibatullin(✉) and Alexander Bogdanov

St. Petersburg University, St Petersburg, Russia
officevg3n@gmail.com, a.v.bogdanov@spbu.ru

Abstract. The increasing complexity of computing and cloud systems, coupled with the growing demand for flexible and efficient solutions, has led to the emergence of efficient management need. This study investigates the potential for improving system manageability through virtualization approaches, with particular emphasis on the integration of systems based on layered abstractions. We examine the scalability challenges present in contemporary systems and argue that layered architecture represents a promising structural approach for addressing these issues. In particular, we focus on the primary limitation of large distributed systems—management complexity. To evaluate this problem, we pose the idea that abstraction layers can simplify system management while the performance overhead introduced by these abstraction mechanisms can be mitigated.

Keywords: Cloud management · Layered systems · Virtualization technologies · Distributed systems · Abstraction overhead

1 Introduction

1.1 Relevance of the Topic

The rapid growth of distributed systems has significantly increased the complexity of system management. This challenge is particularly evident in multi-cloud environments, which we use as a case study in this work, although the conclusions of our research are applicable to distributed systems in general. In such environments, providers expect highly precise and fine-grained management of individual resources. In practice, however, extremely large-scale systems—despite consisting of loosely coupled components—often introduce substantial latency and operational delays. A common approach to addressing the scalability of large systems is to decompose them into smaller, more manageable components. This principle aligns with the concept of *Separation of Concerns* introduced by Dijkstra, E. W. [1]. Also there is an idea of virtual infrastructure [2] that abstracts heterogeneity and simplifies management. Modern system design therefore increasingly adopts microservice or event-driven architectures. High-load systems are rarely implemented as monolithic structures due to their limited scalability and maintainability. Nevertheless, even these architectural approaches exhibit scaling limitations at extreme system sizes. We argue that contemporary

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 481–492, 2027.
https://doi.org/10.1007/978-3-032-30530-5_30

distributed systems require more hybrid architectural models, potentially including adaptive mechanisms capable of evolving in response to emerging operational challenges.

In this paper, we focus on layered architecture based on the following considerations:

1. Advantages:
   (a) **Separation of Concerns**—Each layer is responsible for a specific function. This division reduces system complexity and improves conceptual clarity.
   (b) **Abstraction**—Modifications within one layer typically do not affect other layers as long as the interfaces remain stable. This property allows internal architectures to evolve independently according to the capabilities and requirements of the layer maintainers.
   (c) **Standardization**—Layered architecture is a widely recognized design pattern. Its common structural model facilitates collaboration, maintenance.
2. Limitations:
   (a) **Performance Overhead**—Requests often traverse multiple layers even for simple operations, which can introduce additional latency.
   (b) **Rigid Structure**—Strict separation between layers may hinder cross-layer optimizations or specialized workflows that require tighter integration.
   (c) **Reduced Flexibility**—Operations that deviate from the standard layered flow, such as bypassing layers or integrating external systems, may require additional structural workarounds.

1.2 Literature Review

As noted in the previous section, cloud systems represent a primary example of large-scale distributed systems. This is particularly evident in multicloud environments, federated clouds, and sub-cloud paradigms such as edge computing. To evaluate the relevance of the management problem, we analyzed several recent publications to determine whether this issue is actively addressed in the literature.

Most of the analyzed works are survey papers that focus on current challenges and emerging trends in cloud computing. Figure 1 presents the most frequently mentioned trends identified in these studies, while Fig. 2 summarizes the challenges most commonly discussed by the authors. The analysis indicates that hybrid cloud architectures and resource allocation remain among the most prominent concerns in contemporary research.

We also highlight the study presented in [3], which provides a detailed analysis of potential architectural layers within modern cloud environments. The authors compare several computing paradigms—including fog computing, dew

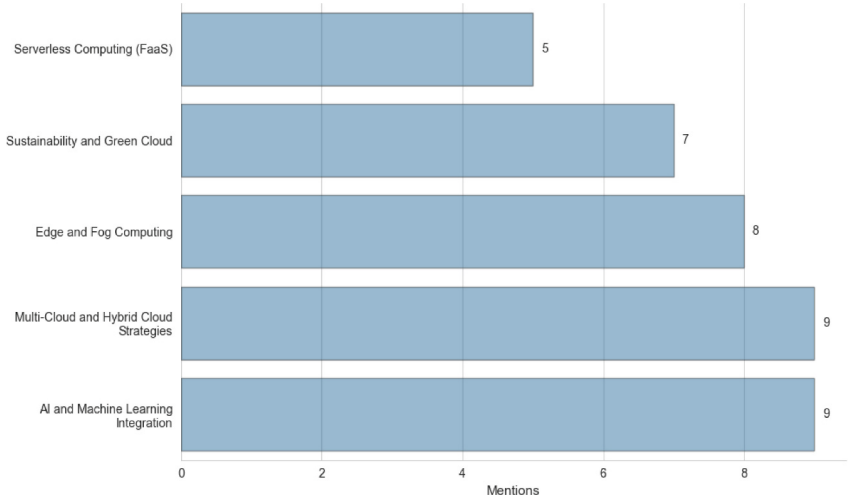


Fig. 1. Trends.

computing, and related models—using evaluation metrics such as latency, security, and manageability. Based on their comparative analysis, we constructed the representation shown in Fig. 3.

This representation highlights an important trade-off: introducing additional architectural layers can reduce end-user latency and improve the proximity of infrastructure to users, but it also increases the complexity of system management. In this context, the trade-off is largely determined by how closely infrastructure components can be positioned relative to end users.

More generally, this trade-off reflects a fundamental property of layered architectures. As systems scale, designers inevitably encounter a balance between improved modularity and the additional coordination required to manage multiple abstraction layers. The objective of this research is to mitigate the management complexity associated with layered architectures, thereby reducing the cost of adopting such scalable system designs.

2 Main Section

2.1 Virtualization Overview

Virtualization denotes an abstraction layer above physical hardware (and sometimes above logical components) that enables treating distributed resources as a cohesive, manageable whole. It can be applied to hardware, OS-level environments, and to interfaces that unify disparate systems.

Primary Technologies

– **Virtual machine (VM):** full emulation of a computing environment (hardware or OS) usually provided by a hypervisor. Strong isolation and OS independence; higher resource footprint and slower startup.

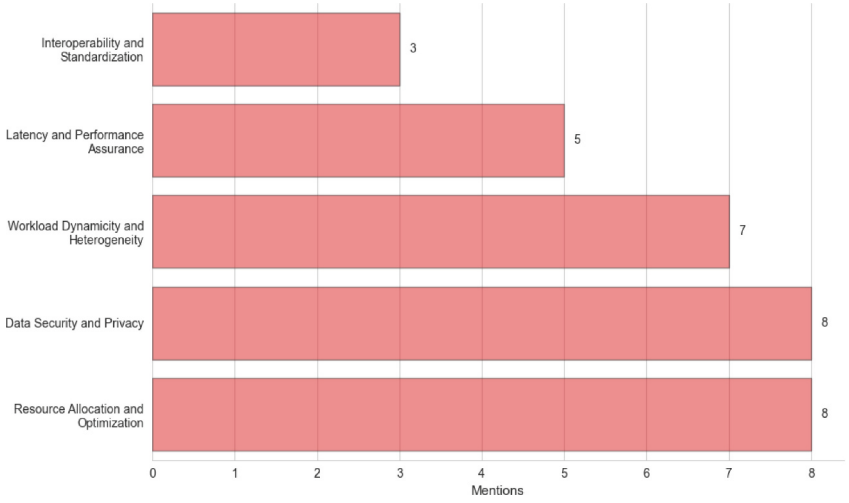


Fig. 2. Challenges.

– **Container:** lightweight isolation of applications using shared-kernel mechanisms. Faster startup and lower overhead than VMs; requires a container runtime.

Containerization and virtual machines have become standard solutions in modern distributed systems. These technologies can also be extended to more complex infrastructures, enabling consistent deployment and management across heterogeneous environments. Research [4] indicates that container-based environments provide an effective balance between flexibility and performance, identifying containers as a particularly suitable choice for distributed system implementations due to their relatively manageable overhead.

Logical Virtualization

– **Adapter/Facade (wrapper):** provides a uniform API over heterogeneous subsystems by adapting or hiding their differences. Useful for integrating components and presenting a stable abstraction to higher layers.
– **Plugin (pluggable interface) architecture:** exposes extension points and well-defined interfaces so independent modules can be added, replaced, or loaded at runtime. Enables extensibility and customization without changing core logic.

This logical or interface-level virtualization is commonly realized with design patterns. By implementing them we hide the system's heterogeneity.

2.2 Hierarchy Overview

Hierarchical models are commonly employed in environments where centralized control over the entire system is infeasible, necessitating delegation of manage-

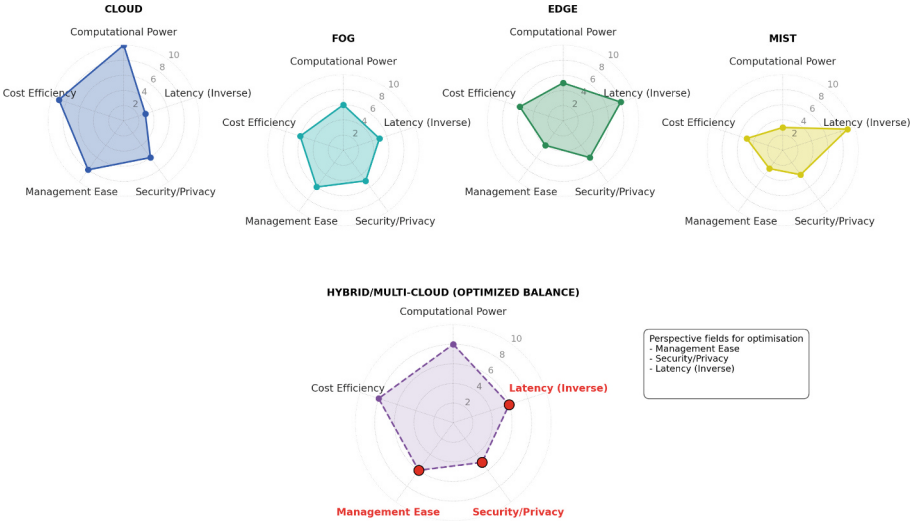


Fig. 3. Latency vs. manageability trade-off.

ment responsibilities to lower layers. Such approaches are typically unnecessary for simple systems and are introduced only when management complexity exceeds practical limits.

This work focuses on large-scale distributed systems as the primary target for abstraction. In general, a hierarchy represents a layered architecture in which higher-level components do not directly control lower-level ones, but instead influence system behavior through defined coordination and control mechanisms to achieve a desired global state.

Common hierarchical structures can be observed in storage systems (e.g., distributed file systems) and in computing architectures (e.g., fast versus slow processing layers). Modern applications, however, are often stateful, requiring a combination of both storage and computational considerations. In contrast, stateless systems primarily depend on computational scalability.

At a fundamental level, layers can be defined in two principal ways:

- **Logical decomposition:** the system is divided into functional blocks. This approach improves transparency and observability, as it clearly identifies which layer (or combination of layers) is responsible for specific operations.
- **Technical decomposition:** the system is partitioned according to technological concerns (e.g., storage, networking, computation). This enhances maintainability by isolating dependencies within well-defined boundaries.

In large-scale distributed systems, an additional criterion becomes critical: **manageability**. When traditional decomposition approaches are insufficient, systems must be structured to minimize management overhead, even at the cost of increased architectural complexity. This perspective assumes that system decomposition is necessary to maintain performance and operability at scale.

A layered system requires at least two levels. A common example is a dual-layer architecture consisting of a fast-access layer and a slower, persistent layer (e.g., in database systems). Upper layers typically provide higher performance at increased cost, while lower layers offer greater capacity and efficiency.

There is no universally optimal number of layers; the appropriate depth depends on system scale, workload characteristics, and management constraints. Excessive layering may introduce overhead, while insufficient layering can limit scalability and control.

2.3 Scaling Limits

In this section, we discuss why scaling in distributed systems has inherent limits. We refer to *Gustafson's law* [5] with communication latency, as modern systems typically evaluate performance and speedup in terms of the increased number of processed jobs.

Additionally, concepts such as *Metcalfe's law* [6] and *Dunbar's number* [7] illustrate the limitations of management and coordination, showing the challenges of scaling system control and the role of separation in overcoming these constraints.

Computation Scaling. The first aspect to consider is computational power. In the ideal case, near-linear scaling can be achieved according to Gustafson's law, provided that the task volume increases proportionally with available processing capacity. However, such conditions are rarely attainable in practice. Communication overhead and resource contention introduce diminishing returns as the system scales. These factors impose practical limits on achievable performance gains, even when computational resources continue to increase.

On Fig. 4 under conditions of efficient task utilization (parallel fraction of 0.3, latency factor of 0.1, and the number of communications approximated as $log_2 n$), a significant performance gap emerges due to latency overhead. From a cost-efficiency perspective, minimizing this gap is critical, as it represents unproductive resource utilization.

In this work, we propose that system decomposition should occur before the cost of such inefficiencies becomes dominant. Scaling within smaller subsystems is more efficient, as task allocation becomes more layer-specific. If computational resources are increasingly consumed by communication overhead, reallocating them toward improving system organization can yield more effective overall performance.

Manageability Scaling. Another key consideration in system decomposition is manageability. There are inherent limits to how much complexity can be effectively managed by humans, both in architectural design and infrastructure operation. This limitation is reflected in Fig. 5 as the blue curve, based on [7].

Automation is typically introduced to extend these limits. However, automation itself becomes insufficient when the system experiences exponential growth

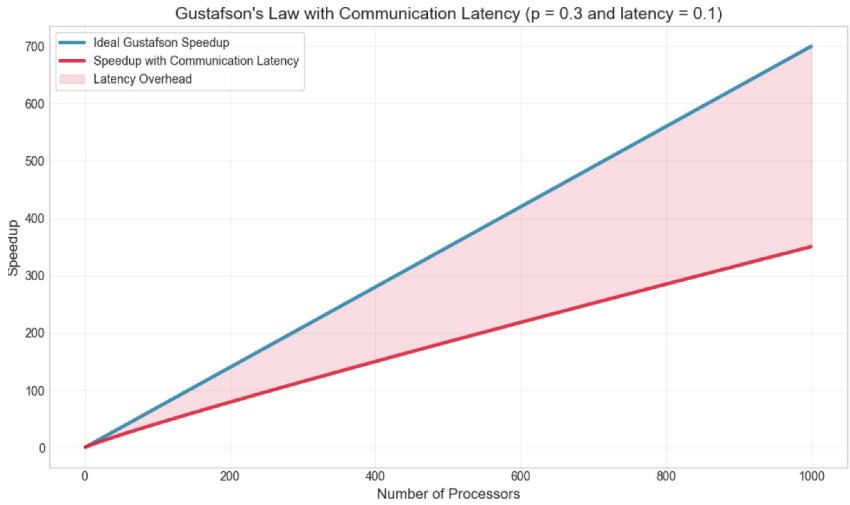


Fig. 4. Gustafson's law with latency.

in the number of nodes or communication channels. This effect is represented by the green curve in Fig. 5.

Therefore, for an efficient *layer split point*, it is necessary to identify the stage at which system manageability remains reliable, but any further increase in complexity would lead to an exponential growth in management difficulty.

2.4 Grid Case Study

A notable example of an intensively distributed system is Grid computing [8], originally developed to support large-scale scientific research. One prominent implementation is deployed at CERN. This system is organized into hierarchical tiers (or layers), each with distinct responsibilities and, in many cases, separate administrative domains. Even within a single tier, management may be distributed among multiple maintainers. Additionally, each layer is based on different technical principles, as illustrated in Fig. 6. This architecture provides a practical example of layered system design in a highly distributed context.

However, Grid systems have not achieved widespread adoption. One of the primary reasons is the complexity introduced by the tiered architecture. As the system scales, challenges arise in coordinating security, resource discovery, scheduling, monitoring, and data management across independent administrative domains. These limitations have contributed to the transition toward cloud computing as a more flexible and manageable paradigm.

Nevertheless, the core idea demonstrated by Grid computing remains relevant. Intensive distribution can be achieved through layered architectures, but modern implementations should leverage advances in cloud, edge, and fog computing to address the limitations observed in Grid systems and enable broader adoption of highly distributed infrastructures.

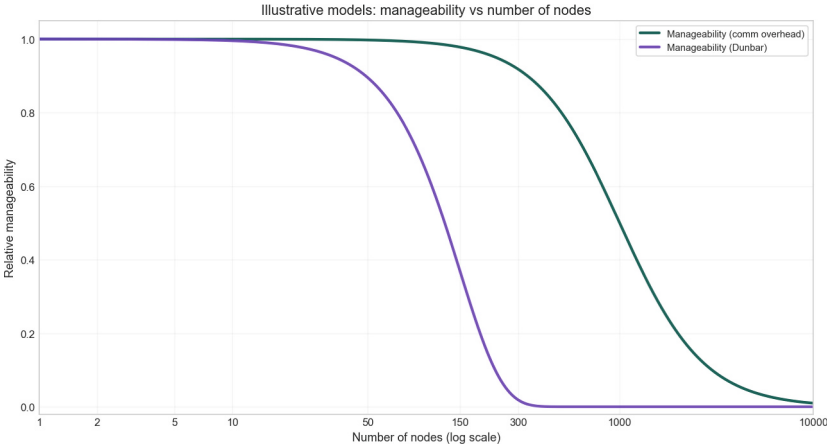


Fig. 5. Manageability.

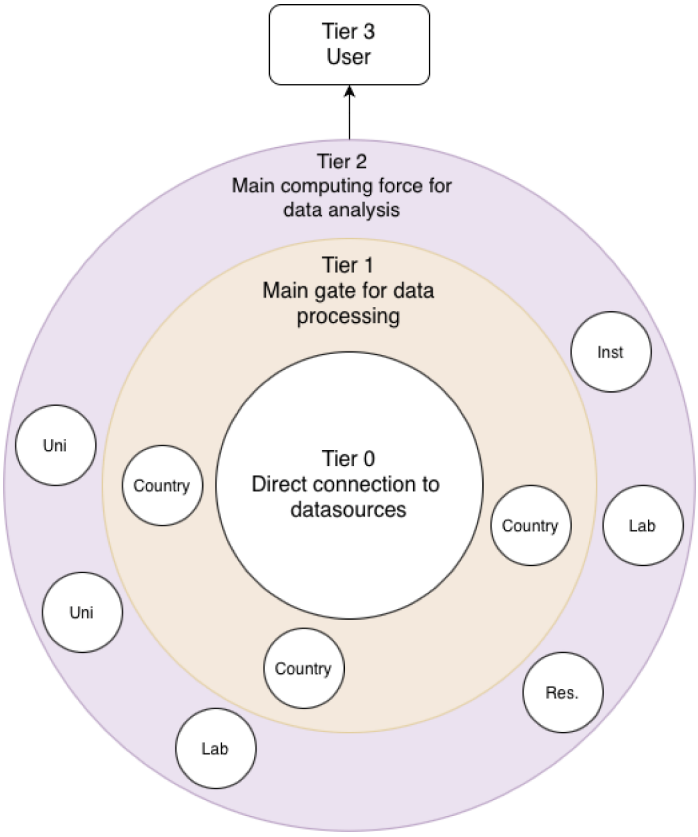


Fig. 6. Infrastructure at CERN.

2.5 Applications

We now outline the conditions under which such a system is expected to be effective, taking into account current technological developments. This approach is not intended for general-purpose distribution tasks. Instead, it is most suitable for systems designed around task- or logic-based distribution, where components and layers are highly decoupled.

Under these conditions, it becomes feasible to identify appropriate *split points* that simplify system management while introducing minimal latency overhead. In the following, we examine three primary problems that are strongly associated with layered architectures:

- AI federative learning
- Data storage systems
- Stateless computational systems

Those are primary cases for effective implementation.

Federated Learning. A current trend [9] in distributed systems is the decentralization of training for deep learning tasks, particularly for large language models. Three primary approaches are commonly used to distribute computation:

1. **Data Parallelism**—replication of the full model across multiple workers.
2. **Model Parallelism**—partitioning of the model across multiple devices.
3. **Federated Learning**—training across decentralized clients that retain local data.

A defining characteristic of federated learning is data heterogeneity. Because client data is decentralized, data samples held by each client may not be independently and identically distributed. Also the access to data is limited by security consideration. Also there could be a case there large datasets are used that could not be moved easily.

This study focuses on federated learning, as it naturally aligns with layered architectures. Such systems can be decomposed into hierarchical levels, typically represented as Cloud → Fog → Edge. On Fig. 7 we give the general idea on how to organize the system based on ≪device reliability≫ as core metrics for manageability.

These layers emerge from the need to process data generated at highly localized devices, where a large number of low-computation nodes operate at the edge. In this structure, management responsibilities are delegated to intermediate layers (e.g., fog), which coordinate aggregation and communication.

While the introduction of intermediate layers adds overhead, directly connecting edge devices to centralized infrastructure leads to increased network ingress, coordination complexity, and synchronization costs that grow with the number of clients.

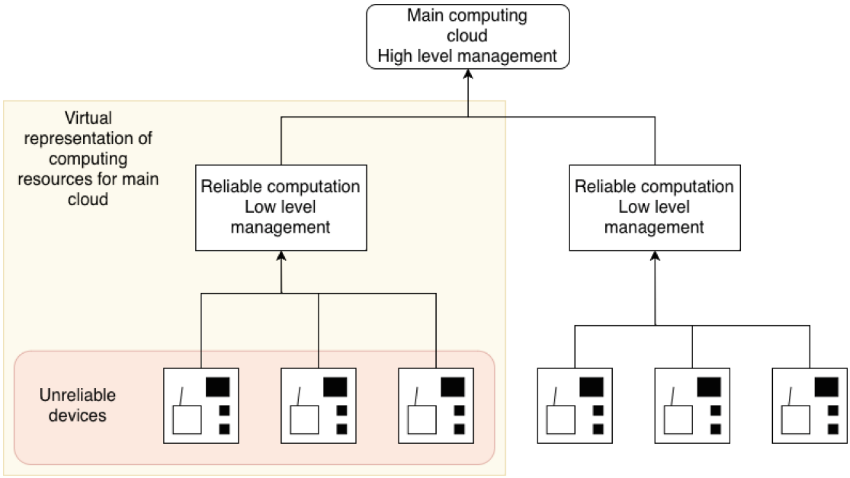


Fig. 7. Federated learning architecture.

Storage Systems. In storage systems, a fundamental trade-off exists between access time and storage cost [10]. On Fig. 8 we demonstrate system separation based on that metrics. Faster storage technologies are significantly more expensive, which leads to the common approach of organizing data into hierarchical layers, such as hot and cold storage. In this case, layered architecture arises naturally, driven primarily by logical and functional partitioning rather than by management scalability.

This approach becomes even more pronounced in the context of large-scale data systems, such as those encountered in Big Data applications. Here, system decomposition is often dictated by data locality. An upper orchestration layer coordinates multiple distributed storage units, each responsible for managing its own subset of data. These units are typically designed to be independent and self-managed, as the high degree of heterogeneity in such environments should remain encapsulated within localized domains.

Both examples demonstrate that, in data-intensive systems, the underlying data characteristics largely determine the system topology. As a result, architectural decisions must be made with careful consideration of these constraints. The most effective scenario occurs when system requirements naturally lead to a layered architecture, allowing this structure to simultaneously address management challenges.

Computational Systems. In computational systems, architectural constraints are less restrictive, as stateless applications can be deployed across a wide range of distributed topologies. However, as the system scales, logical or resource-based partitioning becomes necessary. In such cases, hierarchical structures provide a practical solution, enabling these partitions to improve local manageability while

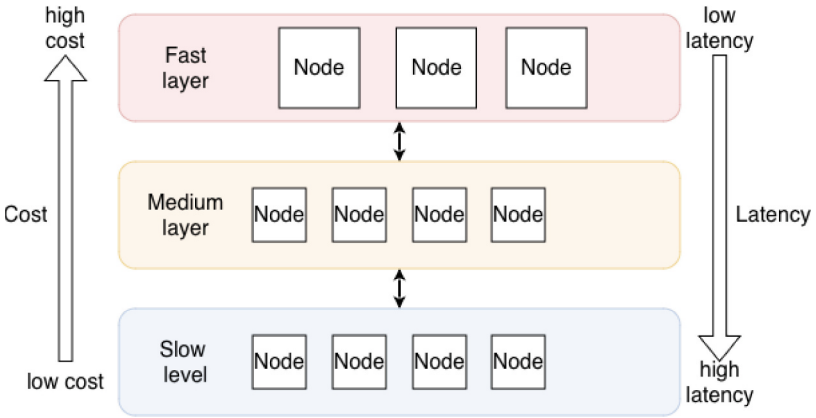


Fig. 8. Latency vs. cost trade-off.

maintaining overall system coherence. This, in turn, shifts the effective scaling limit.

This consideration is particularly relevant for geographically distributed resources. In such environments, a natural approach is to introduce management layers that oversee specific regions and delegate control responsibilities accordingly. This model exemplifies globally distributed systems, where hierarchical organization supports scalability by structuring management across multiple levels.

3 Conclusion

This work examines a potential approach to addressing the increasing complexity of continuously scaling computing systems. We argue that current trends lead to systems that encounter scaling limitations not due to insufficient computational resources, but due to the growing difficulty of managing heterogeneous and highly distributed infrastructures.

We have described the concept of hierarchical architecture as a means of achieving Separation of Concerns in such environments. Conceptual models were presented to highlight the existence of an effective *split point* for system decomposition, where the benefits of layering outweigh the associated overhead. Additionally, we discussed representative cases of highly distributed systems and outlined scenarios in which layered architectures can be effectively applied.

A central recommendation is to identify appropriate metrics that clearly define how a system can be partitioned. Once such metrics are established, they can guide the formation of functional or logical blocks that improve manageability without significantly degrading performance.

The applicability of this architectural approach depends on specific characteristics of real-world systems, many of which are abstracted in this study. Further

empirical evaluation is required for large-scale deployments. Nevertheless, the approach demonstrates potential for extending beyond localized systems and enabling more effective coordination across widely distributed infrastructures.

References

1. Dijkstra, E.W.: On the role of scientific thought. In: Dijkstra, E.W. (ed.) Selected Writings on Computing: A Personal Perspective. Texts and Monographs in Computer Science. Springer, New York (1982). https://doi.org/10.1007/978-1-4612-5695-3_12
2. Korkhov, V., Kobyshev, S., Degtyarev, A., Bogdanov, A.: Light-weight cloud-based virtual computing infrastructure for distributed applications and Hadoop clusters. In: Gervasi, O., et al. (eds.) ICCSA 2017. LNCS, vol. 10408, pp. 399–411. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-62404-4_29
3. Mahapatra, A., Mishra, K., Pradhan, R., et al.: Next generation task offloading techniques in evolving computing paradigms: comparative analysis, current challenges, and future research perspectives. Arch. Comput. Methods Eng. **31**, 1405–1474 (2024)
4. Gankevich, I., et al.: Constructing virtual private supercomputer using virtualization and cloud technologies. In: Murgante, B., et al. (eds.) ICCSA 2014. LNCS, vol. 8584, pp. 341–354. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-09153-2_26
5. Gustafson, J.L.: Reevaluating Amdahl's law. Commun. ACM **31**(5), 532–533 (1988). https://doi.org/10.1145/42411.42415
6. Metcalfe, R.M.: Metcalfe's law: a network becomes more valuable as it reaches more users. Presentation to the 3Com Sales Organization (1983)
7. Dunbar, R.I.M.: Neocortex size as a constraint on group size in primates. J. Hum. Evol. **22**, 469–493 (1992). https://doi.org/10.1016/0047-2484(92)90081-J
8. Foster, I., Kesselman, C.: The History of the Grid. In: High Performance Computing: From Grids and Clouds to Exascale. IOS Press, Amsterdam (2011)
9. Yao, Y.: Federated Large Language Models: Current Progress and Future Directions. arXiv:2409.15723 (2024)
10. Lucas Filho, E.R., Efstathiou, A., Yang, L., et al.: DITIS: an end-to-end system-level simulator and optimizer for distributed tiered storage. SN Comput. Sci. **6**, 746 (2025)

Avg-Frontier: A Low-Overhead Multi-objective Batch Scheduler for Heterogeneous Edge Environments

Aleksei Chernigov and Vladimir Korkhov(✉)

St Petersburg State University, St Petersburg 199034, Russia
v.korkhov@spbu.ru

Abstract. Task scheduling in heterogeneous edge environments involves conflicting objectives related to completion time and energy consumption. Classical low-overhead heuristics usually optimize one dominant criterion, whereas population-based multi-objective methods provide richer trade-off analysis at a higher decision cost. This paper studies a batch scheduler, called *avg_frontier*, intended for repeated scheduling in such environments. The proposed method is organized as a two-stage procedure. First, it constructs an aggregated surrogate model of available devices using average time and energy coefficients over feasible tasks. The scheduler solves a discrete bi-objective allocation problem in the space of task counts, builds a surrogate Pareto frontier, and selects a compromise point according to the *area_ratio* policy controlled by α. Second, the selected allocation is restored into a real task-to-device assignment and re-evaluated by exact simulation. The algorithm is evaluated in SimGrid on two heterogeneous testbeds: a compact three-node profile with 60 tasks and a generated 100-node profile with 6000 tasks. The comparison includes classical heuristics and Pareto-based evolutionary methods. The results show that *avg_frontier* provides a controllable transition between time-oriented and energy-oriented regimes while keeping the scheduler overhead substantially lower than population-based Pareto baselines. The method is not intended to dominate all alternatives in every region of the trade-off, but to provide a computationally cheap mechanism for navigating this trade-off under repeated scheduling.

Keywords: Edge Scheduling · Heterogeneous Systems · Multi-Objective Optimization · Energy-Aware Scheduling · Simulation

1 Introduction

Edge and fog environments are characterized by pronounced hardware and network heterogeneity. In such systems, the scheduler has to operate not only with different computational capabilities of devices, but also with non-uniform communication costs, energy models, and changing availability conditions [14,16,19]. In practice, this leads to a direct conflict between at least two objectives. On

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 493–510, 2027.
https://doi.org/10.1007/978-3-032-30530-5_31

the one hand, the system must process tasks with low completion time. On the other hand, the same decisions influence energy consumption of both computing devices and communication channels.

A significant part of the scheduling literature relies on classical heuristics. These methods remain relevant because they are computationally simple and often provide reasonable schedules in heterogeneous systems [1,11–13,15,22,24]. However, such methods are usually built around a single dominant criterion or around a fixed deterministic rule. As a consequence, they are less convenient in the case when one needs to move between time-oriented and energy-oriented regimes without changing the scheduling framework itself.

Another direction is the use of multi-objective optimization methods, including evolutionary algorithms and decomposition-based approaches [6,7,23,25,26]. These methods are natural for approximating the Pareto set. At the same time, their use inside an operational scheduler raises a practical question: if the scheduling decision is repeated many times during system execution, the decision procedure itself becomes part of the total system cost.

The purpose of this paper is to study a lightweight multi-objective batch scheduler for heterogeneous edge environments. The scheduler, called *avg_frontier*, is based on a different principle than direct item-level Pareto optimization. Instead of searching in the full combinatorial space of task-to-device assignments, it first compresses the current batch into a surrogate allocation problem defined over the numbers of tasks assigned to devices. In this aggregated space, the method constructs an explicit discrete Pareto frontier with respect to surrogate makespan and surrogate active energy, selects one compromise point according to a prescribed front policy, and only after that restores the corresponding item-level assignment. Thus, the multi-objective search is performed in a substantially smaller space, while the final plan remains connected to the exact simulated execution model.

The contribution of the paper can be summarized as follows.

1. A compact mathematical formulation of the batch scheduling problem is given together with its surrogate aggregated version in the space of device load counts, where makespan and active energy remain explicit optimization criteria.
2. A low-overhead scheduler, *avg_frontier*, is described. The method explicitly builds a discrete surrogate Pareto frontier, selects one operating point using the *area_ratio* policy controlled by parameter α, restores the corresponding item-level assignment, and then performs exact re-evaluation of the obtained plan.
3. A simulation-based evaluation is performed on two heterogeneous testbeds: a compact three-node profile and a generated 100-node profile. The comparison includes classical heuristics and Pareto-based evolutionary algorithms and explicitly considers both schedule quality and scheduler overhead.

The rest of the paper is organized as follows. Section 2 summarizes related work. Section 3 introduces the problem formulation. Section 4 describes the pro-

posed scheduler. Section 5 presents the simulation environment and the experimental profiles. Section 6 discusses the obtained results. Section 7 outlines the main limitations of the study. Section 8 concludes the paper.

2 Related Work

2.1 Scheduling in Heterogeneous Environments

Task mapping in heterogeneous systems has been extensively studied. Surveys and comparative studies cover both static scheduling and independent-task mapping under common assumptions [1,12]. Classical methods such as OLB, MET, MCT, Min-Min, Max-Min, and Sufferage remain widely used baselines because they are simple and have low decision cost [13,15]. For dependent tasks, HEFT remains a standard example of low-complexity list scheduling [22].

These heuristics are attractive in practice, but they are usually centered on one dominant criterion, most often completion time. This makes them less convenient when the scheduler must move between time-oriented and energy-oriented regimes within one framework.

2.2 Energy-Aware and Multi-objective Scheduling

Joint consideration of time and energy has led to energy-aware scheduling methods for heterogeneous and distributed systems [4,10,17]. Specialized examples include energy-efficient scheduling for batch-of-tasks applications and weighted energy- and makespan-aware scheduling in cloud environments [18,21].

Multi-objective optimization provides a natural framework for such problems [6]. Frequently used methods include NSGA-II, SPEA2, MOEA/D, and multi-objective local search [7,9,23,26]. These methods approximate the Pareto set, but their decision cost becomes important when optimization is repeated during normal scheduler operation. This creates an intermediate niche between lightweight single-solution heuristics and more expensive Pareto-based procedures.

Thus, a gap remains between two classes of approaches. On one side, there are lightweight heuristics with limited multi-objective flexibility. On the other side, there are more expressive Pareto-based procedures with a substantially larger decision cost. The present work addresses this intermediate space.

2.3 Edge Orchestration and Simulation Environments

Edge computing has intensified interest in resource management, placement, and orchestration under heterogeneous constraints [14,16,19]. Examples include fog application deployment, fuzzy workload orchestration, and multi-objective provisioning in edge clouds [2,5,20]. These studies emphasize that operational decisions must often be made repeatedly under changing conditions.

Simulation is useful here because real edge infrastructures are costly to control and reproduce. iFogSim is a common tool in this area [8]. In this work,

SimGrid is used because it provides a scalable basis for modeling distributed platforms [3]. This allows explicit representation of device heterogeneity, communication parameters, and energy-aware metrics while preserving repeatability of the experiments.

3 Problem Formulation

Let

$$J = \{1, \ldots, N\}$$

be the current batch of independent tasks and

$$I = \{1, \ldots, M\}$$

be the set of available compute devices.

Each task $j \in J$ is characterized by its input data size, computational demand, output data size, and source node. In the present study, the scalar quantity

$$w_j = \max(0, \text{computeSize}_j)$$

is used as the main measure of task computational complexity.

A scheduling decision is defined as a mapping

$$\phi : J \to I,$$

where $\phi(j) = i$ means that task j is assigned to device i. In the present model, the communication route is not an independent decision variable. For every task–device pair (j, i), the route used for input and output transfer is determined by the routing policy and the current system state. Therefore, feasibility and the corresponding time and energy estimates are defined for the induced task–device assignment rather than optimized over routes separately.

For each feasible pair (j, i), the simulator provides route-aware estimates of input transfer time, execution time, and output transfer time,

$$t_{ij}^{in}, \qquad t_{ij}^{ex}, \qquad t_{ij}^{out},$$

as well as the corresponding energy estimates,

$$e_{ij}^{in}, \qquad e_{ij}^{ex}, \qquad e_{ij}^{out}.$$

Accordingly, the estimated completion-time contribution of assigning task j to device i is

$$T_{ij} = t_{ij}^{in} + t_{ij}^{ex} + t_{ij}^{out},$$

and the corresponding active-energy contribution is

$$E_{ij} = e_{ij}^{in} + e_{ij}^{ex} + e_{ij}^{out}.$$

Since devices may already be partially occupied at the scheduling instant, the final completion time of the whole batch depends not only on the task-to-device mapping itself, but also on the current device loads, queueing effects, and the order in which assigned tasks are executed. Therefore, the exact batch makespan is a functional of the full assignment ϕ and of the current system state. We denote it by

$$T(\phi).$$

Similarly, the exact active energy of the obtained plan is denoted by

$$E(\phi).$$

The scheduling problem considered in this paper is the following bi-objective optimization problem:

$$\min_{\phi} \big(T(\phi), E(\phi)\big),$$

subject to feasibility of all assignments contained in ϕ.

Thus, the scheduler must assign all tasks in the current batch to heterogeneous devices while balancing completion time and active energy. Direct optimization in the full assignment space is combinatorial and becomes impractical when the decision is rebuilt repeatedly. Therefore, the proposed method uses a lower-dimensional surrogate problem, restores its solution into an item-level assignment, and re-evaluates the result by exact simulation.

4 Avg-Frontier Scheduling Algorithm

The proposed scheduler solves the original batch assignment problem indirectly. Its construction consists of two stages. At the first stage, the current batch is replaced by an aggregated surrogate allocation problem. Instead of assigning each task to a concrete device immediately, the scheduler first determines only the number of tasks that should be sent to each device. This reduced problem is used to construct a discrete surrogate Pareto frontier with respect to makespan and active energy. At the second stage, the selected surrogate allocation is restored into a concrete item-level assignment, after which the obtained plan is re-evaluated by exact simulation. Thus, the main idea of *avg_frontier* is to perform multi-objective search in a lower-dimensional aggregated space and only then reconstruct a feasible detailed schedule. The overall workflow of the proposed method is illustrated in Fig. 1.

Let $F_i \subseteq J$ be the set of tasks feasible for device i in the current system state. For every feasible pair (i, j), the simulator provides route-aware estimates of input, execution, and output times

$$t_{ij}^{in}, \quad t_{ij}^{ex}, \quad t_{ij}^{out},$$

and the corresponding energy estimates

$$e_{ij}^{in}, \quad e_{ij}^{ex}, \quad e_{ij}^{out}.$$

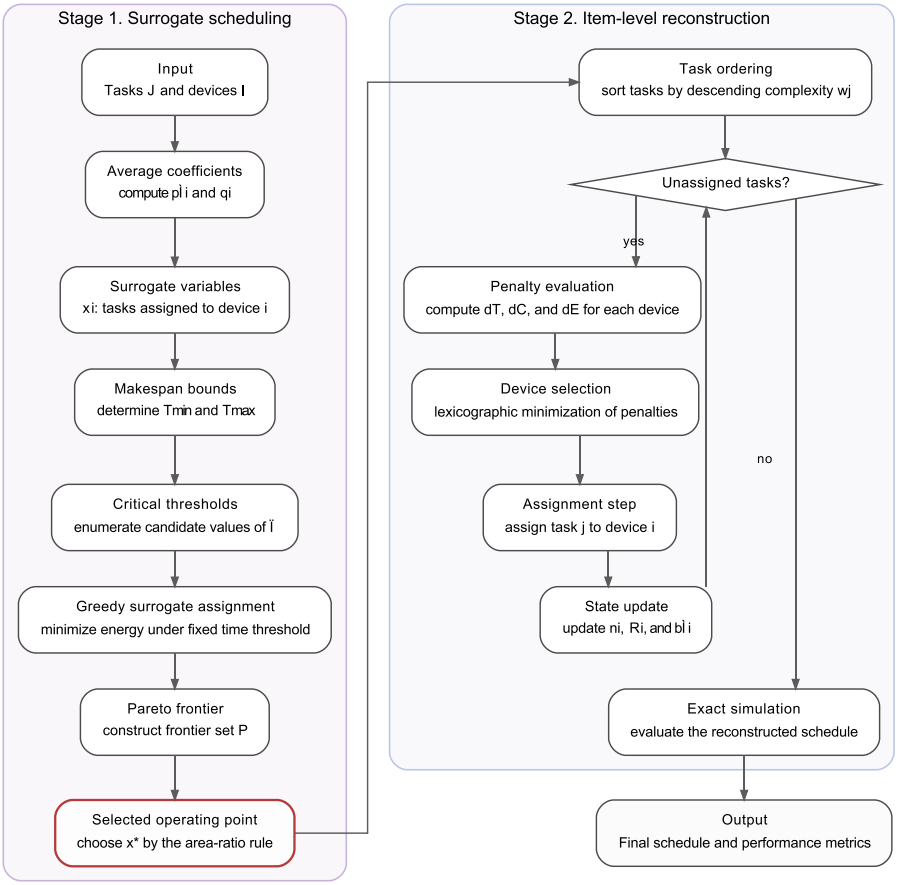


Fig. 1. Pipeline of the *avg_frontier* scheduling algorithm.

Using the task complexity measure w_j, define

$$W_i^{feas} = \sum_{j \in F_i} w_j.$$

Then the average time coefficients per unit complexity are

$$p_i^{in} = \frac{\sum_{j \in F_i} t_{ij}^{in}}{W_i^{feas}}, \qquad p_i^{ex} = \frac{\sum_{j \in F_i} t_{ij}^{ex}}{W_i^{feas}}, \qquad p_i^{out} = \frac{\sum_{j \in F_i} t_{ij}^{out}}{W_i^{feas}},$$

and the average energy coefficients are

$$q_i^{in} = \frac{\sum_{j \in F_i} e_{ij}^{in}}{W_i^{feas}}, \qquad q_i^{ex} = \frac{\sum_{j \in F_i} e_{ij}^{ex}}{W_i^{feas}}, \qquad q_i^{out} = \frac{\sum_{j \in F_i} e_{ij}^{out}}{W_i^{feas}}.$$

The total specific active-energy coefficient is

$$q_i = q_i^{in} + q_i^{ex} + q_i^{out}.$$

To approximate the time contribution of device i, the scheduler uses

$$\hat{p}_i = p_i^{ex} + \beta_{in} p_i^{in} + \beta_{out} p_i^{out},$$

where $\beta_{in} \geq 0$ and $\beta_{out} \geq 0$ control the influence of communication on the surrogate time model. In the reported experiments we use $\beta_{in} = 1$ and $\beta_{out} = 0$. This choice reflects the execution model adopted in the simulator.

In addition, each device is characterized by its current ready time b_i and by the maximum number of currently feasible tasks

$$u_i = |F_i|.$$

Let

$$W = \sum_{j=1}^{N} w_j, \qquad \bar{c} = \frac{W}{N}$$

be the total and average complexity of the current batch. The surrogate decision variable is the integer vector

$$x = (x_1, \ldots, x_M),$$

where x_i denotes the number of tasks assigned to device i. It satisfies

$$\sum_{i=1}^{M} x_i = N, \qquad 0 \leq x_i \leq u_i.$$

Under the average-complexity approximation, the surrogate completion time of device i is

$$T_i(x_i) = b_i + \hat{p}_i \bar{c}\, x_i.$$

Hence, the surrogate makespan and active energy are defined as

$$T(x) = \max_{i \in I} \left(b_i + \hat{p}_i \bar{c}\, x_i \right), \qquad E(x) = \sum_{i=1}^{M} q_i \bar{c}\, x_i.$$

Thus, the first stage solves the discrete bi-objective problem

$$\min_x \big(T(x), E(x) \big)$$

subject to the above feasibility constraints.

To construct the surrogate frontier, for a fixed time threshold τ define the capacity of device i as

$$\mathrm{cap}_i(\tau) = \max\left(0, \min\left\{ N,\ u_i,\ \left\lfloor \frac{\tau - b_i}{\hat{p}_i \bar{c}} \right\rfloor \right\} \right).$$

The value τ is feasible if

$$\sum_{i=1}^{M} \mathrm{cap}_i(\tau) \geq N.$$

Therefore, the minimum achievable surrogate makespan is

$$T_{\min} = \min \left\{ \tau : \sum_{i=1}^{M} \mathrm{cap}_i(\tau) \geq N \right\}.$$

In practice, it is sufficient to enumerate only critical values

$$\tau = b_i + \hat{p}_i \bar{c}\, r, \qquad r = 0, 1, \ldots, \min(N, u_i),$$

because the capacities change only at these points.

For each feasible τ, the corresponding minimum-energy allocation is obtained by minimizing

$$E(x) = \sum_{i=1}^{M} q_i \bar{c}\, x_i$$

under the constraints

$$\sum_{i=1}^{M} x_i = N, \qquad 0 \leq x_i \leq \mathrm{cap}_i(\tau).$$

Since the objective is linear, this allocation is constructed greedily: devices are sorted by nondecreasing $(q_i, \hat{p}_i, b_i, i)$, and tasks are assigned sequentially until all N tasks are distributed. Separately, the scheduler constructs the energy-extreme feasible allocation

$$x^E = \arg\min E(x)$$

subject to

$$\sum_{i=1}^{M} x_i = N, \qquad 0 \leq x_i \leq u_i.$$

That is, x^E is the feasible surrogate allocation with minimum active energy among all allocations assigning all N tasks. Since the objective is linear, x^E is obtained greedily by sorting devices by nondecreasing $(q_i, \hat{p}_i, b_i, i)$ and filling their capacities until all N tasks are assigned. We then define

$$T_{\max} = T(x^E).$$

Then all critical values $\tau \in [T_{\min}, T_{\max}]$ are enumerated. For each such τ, the corresponding energy-minimizing allocation $x(\tau)$ is constructed, and the point

$$\big(T(x(\tau)), E(x(\tau))\big)$$

is added to the candidate set. After dominated points are removed, the remaining set

$$P = \{x^{(1)}, x^{(2)}, \ldots, x^{(K)}\}$$

forms the discrete surrogate Pareto frontier.

One frontier point is then selected using the *area_ ratio* rule. Assume that the frontier points are ordered by increasing surrogate makespan. Define the ideal corner

$$(T^{\dagger}, E^{\dagger}) = \big(T(x^{(1)}), E(x^{(K)})\big).$$

For each adjacent pair of frontier points, compute the triangle area

$$A_k = \mathrm{Area}\Big((T^{\dagger}, E^{\dagger}), (T(x^{(k)}), E(x^{(k)})), (T(x^{(k+1)}), E(x^{(k+1)}))\Big), \qquad k = 1, \ldots, K-1,$$

and then the cumulative area ratio

$$\rho_k = \frac{\sum_{s=1}^{k-1} A_s}{\sum_{s=1}^{K-1} A_s}, \qquad k = 1, \ldots, K.$$

The selected surrogate point is

$$x^{\star} = \arg\min_{x^{(k)} \in P} |\rho_k - \alpha|,$$

where $\alpha \in [0, 1]$ is the preference parameter. Small α values correspond to the time-oriented part of the frontier, while large values correspond to the energy-oriented part.

After the surrogate point

$$x^{\star} = (x_1^{\star}, \ldots, x_M^{\star})$$

has been selected, the scheduler restores a real item-level assignment. For each device i, the initial restoration targets are

$$n_i^{(0)} = x_i^{\star}, \qquad R_i^{(0)} = \bar{c} x_i^{\star}, \qquad \tilde{b}_i^{(0)} = b_i,$$

where n_i is the remaining number of tasks to assign, R_i is the remaining target complexity, and $\tilde{b}_i$ is the current virtual surrogate load. Tasks are processed in nonincreasing order of complexity in order to preserve the selected surrogate structure more accurately.

For the current task j with complexity w_j, each feasible device i with $n_i > 0$ is evaluated. The updated virtual load is

$$\tilde{b}_i^{+} = \tilde{b}_i + \hat{p}_i w_j,$$

and the corresponding time penalty is

$$\Delta T_i = \max\{0, \tilde{b}_i^{+} - T(x^{\star})\}.$$

The structural state is updated as

$$R_i^{+} = R_i - w_j, \qquad n_i^{+} = n_i - 1.$$

If the remaining unscheduled tasks have total complexity W_{rem} and count N_{rem}, define

$$\bar{w}_{rem} = \begin{cases} W_{rem}/N_{rem}, & N_{rem} > 0, \\ 0, & N_{rem} = 0. \end{cases}$$

Then the structural deviation penalty is

$$\Delta C_i = \begin{cases} |R_i^+|, & n_i^+ = 0, \\ \left| \dfrac{R_i^+}{n_i^+} - \bar{w}_{rem} \right|, & n_i^+ > 0. \end{cases}$$

If $R_i^+ < 0$, the penalty is additionally increased by $|R_i^+|$. The approximate energy increment is

$$\Delta E_i = q_i w_j.$$

The selected device is the one that lexicographically minimizes

$$(\Delta T_i, \Delta C_i, \Delta E_i, \tau_i, \pi_i, \nu_i),$$

where τ_i is the exact completion-time estimate of the candidate assignment, π_i is a deterministic route-based tie-break, and ν_i is the node identifier. After assigning the task, the corresponding values of n_i, R_i, and $\tilde{b}_i$ are updated. If no device with $n_i > 0$ yields a feasible candidate, a fallback rule is used: the task is assigned to the feasible device with the minimum exact expected completion time.

Finally, after the full item-level assignment has been restored, the scheduler performs exact simulation of the obtained fixed plan. Let

$$\phi^\star : J \to I$$

be the reconstructed mapping. The simulator evaluates its exact characteristics

$$\big(T_{exact}(\phi^\star), E_{exact}(\phi^\star)\big),$$

and these exact values replace the surrogate estimates in the final stored decision. Therefore, the surrogate stage is used only to construct a compact search space and to select a compromise operating point, whereas the final schedule is always validated by the exact execution model.

5 Simulation Environment

5.1 General Setup

The experiments were conducted in a SimGrid-based discrete-event environment [3]. Platform descriptions included an XML topology, CSV files for device and channel types, and CSV files for workload and scenario description. This representation allowed the experimental profile to be changed without modifying the scheduling code.

Each compute node had a device type describing performance, idle and busy power, and transfer capabilities. Channels were described by bandwidth, latency, transmission and reception power, and startup and tail effects. The scheduler was invoked in *batch* mode: at each invocation it observed the current state, scheduled the newly available batch, and the simulation then proceeded to execution.

5.2 Workload Design

One task family was used to keep the comparison focused on scheduling behavior rather than on interactions between workload classes. The tasks were independent and batch-generated; for each task, the simulator modeled input transfer, computation, and output transfer. The small testbed served as an interpretable time–energy case, while the large testbed was used to evaluate decision cost under more devices and repeated scheduling calls.

5.3 Experimental Setup

Two testbeds were used. The small testbed depicted in Fig. 2, contained three heterogeneous compute nodes, $N0$, $N1$, and $N2$, with different performance, power, and communication characteristics. Its communication layer included WiFi-, USB-, and cellular-type links.

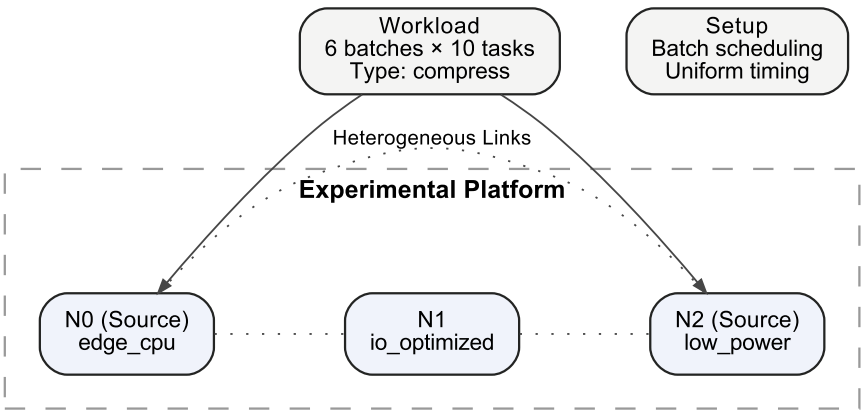


Fig. 2. Small experimental profile used in the simulation study. The platform contains three heterogeneous compute nodes connected by heterogeneous communication links. The workload consists of six consecutive batches of independent *compress* tasks generated from the available source nodes.

The small workload consisted of six batches of 10 independent *compress* tasks, for a total of 60 tasks, generated from $N0$, $N1$, and $N2$.

The large testbed was a generated 100-node profile. Node types were assigned cyclically from three classes: *edge_ cpu*, *io_ optimized*, and *low_ power*. The topology used offset-based connectivity; each logical edge had WiFi-, USB-, and cellular-type links. Five source nodes were placed approximately uniformly: $N0$, $N25$, $N50$, $N74$, and $N99$. The workload contained 600 batches of 10 independent tasks, for a total of 6000 tasks.

5.4 Baselines and Metrics

To evaluate the proposed method, it was compared with several groups of baseline schedulers.

The comparison included three groups of schedulers. The first group contained classical heuristics: `source_only`, `rr`, `olb`, `met`, `mct`, `min_min`, `max_min`, and `sufferage`. The second group contained energy-aware or hybrid heuristics: `minmin_energy`, `minmin_energy_time`, `minsuff`, `lws_greedy`, and `g2`. The third group contained Pareto-based methods: `nsga2`, `moead`, `spea2`, and `me_mls`. The `me_mls` method was evaluated only on the small testbed because it did not complete on the large testbed within practical time.

Implementation parameters were fixed as follows. For NSGA-II, MOEA/D, SPEA2, and ME-MLS, when used, the population size was 64, the archive size was 64, and the number of main iterations/generations was 60. For MOEA/D, the neighborhood size was 10 and the replacement limit was 2. For ME-MLS, the number of local-search iterations was 200. The crossover and mutation probabilities were set to 0.90 and 0.10, respectively. Seed solution generation used $\alpha_{seed} = 0.5$ and $\beta_{seed} = 0.5$. All experiments were run with a fixed random seed equal to 42. Representative points reported for multi-objective baselines were selected from the obtained nondominated sets using the same *area_ ratio* rule as for *avg_ frontier*. For `lws_greedy` and `g2`, equal weights were used for time and energy components (0.5/0.5). For *avg_ frontier*, the surrogate communication weights were fixed to $\beta_{in} = 1$ and $\beta_{out} = 0$, and tasks were restored in nonincreasing order of complexity.

For each method, the following metrics were collected:

- *successRate*, that is, the fraction of tasks completed successfully;
- *makespan*, which is the time it takes to complete all tasks;
- *activeEnergyJ*, that is, the active energy consumption measured in joules;
- *schedulerRealWallTimeSec*, that is, the real wall-clock time required for all scheduling decisions.

Among these metrics, makespan and active energy correspond directly to the optimization problem considered in this work. The scheduler wall-clock time was additionally recorded in order to estimate the practical computational cost of the decision procedure.

6 Results

6.1 General Remarks on Result Interpretation

All completed runs in both testbeds achieved $successRate = 1.0$, which allows direct comparison of the collected time and energy metrics without additional penalty handling. The reported comparison is based on makespan, active energy, and scheduler wall-clock time.

The two testbeds play different roles in the analysis. The three-node profile is used as a compact interpretable case in which the assignment structure and the time–energy trade-off can be inspected directly. The 100-node profile is used as a larger stress case for repeated scheduling and for comparing the decision cost of lightweight and population-based methods. Therefore, the conclusions should be interpreted as a controlled comparison on two simulated configurations rather than as an exhaustive empirical characterization of all possible edge platforms.

6.2 Overall Behavior of *avg_frontier*

The results show that *avg_frontier* provides a controllable shift from time-oriented to energy-oriented operating modes as α increases from 0.0 to 1.0, as visualized in Fig. 3. On the small testbed, its makespan increases from 438.40 s to 1243.39 s, while active energy decreases from 64.02 kJ to 39.35 kJ. On the large testbed, the same change in α moves the method from a time-oriented point with makespan 1931.43 s and active energy 23.99 MJ to an energy-oriented point with makespan 40328.86 s and active energy 4.06 MJ.

This behavior confirms the intended role of α: small values select operating points closer to the time-oriented region, whereas large values shift the allocation toward energy-saving devices. The transition is not meant to produce the best value of every metric simultaneously. Rather, it provides a single parameter for navigating the time–energy trade-off.

The scheduler overhead also remains substantially lower than that of population-based Pareto baselines. Although the absolute wall-clock time of *avg_frontier* increases on the 100-node testbed, it stays in the range of lightweight batch heuristics and remains far below the evolutionary methods, especially under repeated scheduling.

6.3 Comparison with Classical Baselines

On the small testbed, the time-oriented setting of *avg_frontier* is stronger than the classical batch heuristics in terms of makespan. At $\alpha = 0.0$, it achieves a makespan of 438.40 s, compared with 456.14 s for `sufferage`, 456.63 s for `max_min`, and 495.61 s for `mct`. At the energy-oriented end, *avg_frontier* approaches the energy-focused heuristics: at $\alpha = 1.0$, its active energy is 39.35 kJ, while `minsuff`, `minmin_energy`, and `minmin_energy_time` obtain 37.71–40.56 kJ.

On the large testbed, the picture is more nuanced. Simple load-distribution heuristics such as `olb` and `mct` remain strong in the time-oriented region, whereas

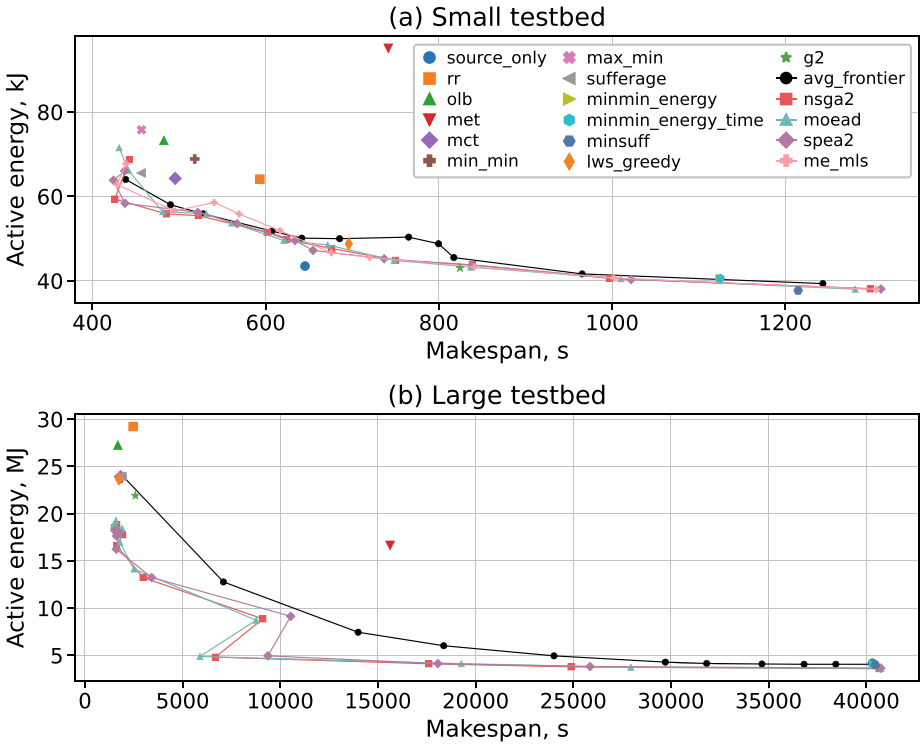


Fig. 3. Comparison of scheduling results in the makespan–active energy plane for the small and large testbeds. Lines correspond to methods producing multiple trade-off solutions, while isolated markers denote single-solution baselines. Energy is shown in kJ for the small testbed and in MJ for the large testbed.

explicitly energy-oriented heuristics obtain the lowest active energy among classical methods at the cost of very large makespan. The *avg_frontier* curve connects these regimes within one framework: as α grows, the method moves from a competitive time-oriented allocation toward energy-oriented behavior close to the energy-focused baselines.

Therefore, *avg_frontier* should be interpreted not as a method that simultaneously minimizes all metrics, but as a lightweight mechanism for controlled movement along the time–energy trade-off.

6.4 Comparison with Pareto-Based Baselines

The Pareto-based baselines also produce trade-off curves in the makespan–energy plane. In several regions, especially on the large testbed, they can find better schedule-quality points than *avg_frontier*. For example, in the time-oriented region of the large testbed, NSGA-II, MOEA/D, and SPEA2 reach

lower makespan values than *avg_frontier*. This is expected because these methods search directly in a richer assignment space.

However, the main practical difference is the cost of the decision procedure itself. Figure 4 summarizes this difference on both testbeds.

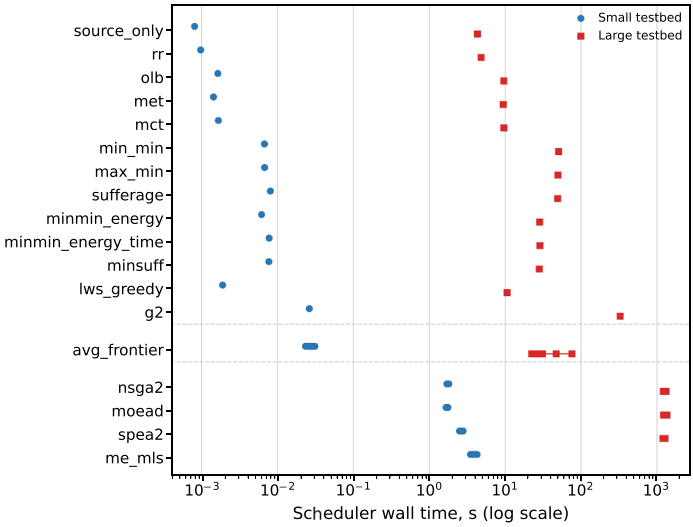


Fig. 4. Scheduler wall-clock time on the small and large testbeds. The horizontal axis uses a logarithmic scale. For methods producing several operating points, markers correspond to different α settings, while horizontal segments indicate the observed range of scheduler times.

On the small testbed, *avg_frontier* requires only a small fraction of the wall-clock time required by the population-based baselines. On the large testbed, this difference becomes much more pronounced: NSGA-II, MOEA/D, and SPEA2 require scheduler wall-clock times on the order of 10^3 seconds, whereas *avg_frontier* remains on the order of tens of seconds over all tested α values. ME-MLS was not included in the large-testbed comparison because it did not complete within practical time in preliminary runs.

Thus, the main advantage of *avg_frontier* is not unconditional dominance in the makespan–energy plane. Its advantage is the balance between schedule quality, controllability of the operating point, and decision cost. This balance is especially important when scheduling is invoked repeatedly and the scheduler itself becomes part of the system overhead.

7 Limitations

The present study has several limitations that should be taken into account when interpreting the results.

First, the evaluation is restricted to two simulated testbeds and one task family. The small testbed was designed for interpretability, whereas the large testbed was generated to study scalability and scheduler overhead. These configurations cover two different scales, but they still do not provide an exhaustive characterization over multiple workload classes, topology types, or network/compute ratios.

Second, the evaluation is simulation-based. This is appropriate for controlled comparative analysis, but practical deployment will additionally depend on aspects that are harder to capture in the present model, including operating-system overheads, measurement noise, and implementation details of the runtime layer.

Third, the surrogate frontier should not be interpreted as an exact approximation of the true global Pareto front of the original assignment problem. It is used as a compact decision mechanism whose selected allocations are then re-evaluated by exact simulation. The reported results therefore characterize the behavior of this mechanism rather than the full global Pareto structure of the scheduling problem.

These limitations do not invalidate the obtained observations, but they delineate the scope within which the conclusions of the paper should be understood.

8 Conclusion

This paper considered the problem of batch scheduling of independent tasks in a heterogeneous edge environment under conflicting time and energy requirements. A lightweight scheduler, *avg_frontier*, was studied. The method first solves an aggregated surrogate bi-objective allocation problem in the space of task counts assigned to devices, explicitly constructs the corresponding discrete Pareto frontier, selects one operating point by the *area_ratio* policy, restores the resulting item-level assignment, and then re-evaluates the obtained plan by exact simulation.

The experiments on the small and large heterogeneous testbeds lead to the following conclusions. First, the proposed scheduler provides a clear and controllable transition between more time-oriented and more energy-oriented operating modes through a single parameter α. Second, its scheduler overhead remains substantially lower than that of population-based Pareto baselines, and this difference becomes especially visible on the 100-node testbed. Third, although *avg_frontier* does not dominate all alternatives in the makespan–energy plane, it provides a favorable compromise between schedule quality, interpretability of the selected operating point, and decision cost.

At the same time, the present conclusions should be interpreted carefully. The evaluation remains limited to simulated configurations and one task family. Therefore, the main result of the paper is not a universal claim that *avg_frontier* dominates all alternatives, but the demonstration that this scheduler provides an interpretable and computationally cheap mechanism for navigating the time–energy trade-off under repeated scheduling.

A natural continuation of this work is associated with three directions: validation on additional platform and workload configurations, investigation of adaptive policies for selecting the target operating regime, and comparison of simulation-based predictions with execution traces from a physical prototype.

References

1. Braun, T.D., et al.: A comparison of eleven static heuristics for mapping a class of independent tasks onto heterogeneous distributed computing systems. J. Parallel Distrib. Comput. **61**(6), 810–837 (2001)
2. Brogi, A., Forti, S., Ibrahim, A.: How to best deploy your fog applications, probably. In: IEEE 1st International Conference on Fog and Edge Computing (ICFEC), pp. 105–114. IEEE (2017)
3. Casanova, H., Giersch, A., Legrand, A., Quinson, M., Suter, F.: Versatile, scalable, and accurate simulation of distributed applications and platforms. J. Parallel Distrib. Comput. **74**(10), 2899–2917 (2014)
4. Czarnul, P., Proficz, J., Krzywaniak, A.: Energy-aware high-performance computing: survey of state-of-the-art tools, techniques, and environments. Sci. Program. **2019**(1), 8348791 (2019)
5. De Maio, V., Brandic, I.: Multi-objective mobile edge provisioning in small cell clouds. In: Proceedings of the 2019 ACM/SPEC International Conference on Performance Engineering, pp. 127–138 (2019)
6. Deb, K.: Multi-objective optimisation using evolutionary algorithms: an introduction. In: Wang, L., Ng, A., Deb, K. (eds.) Multi-Objective Evolutionary Optimisation for Product Design and Manufacturing, pp. 3–34. Springer, London (2011). https://doi.org/10.1007/978-0-85729-652-8_1
7. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comput. **6**(2), 182–197 (2002)
8. Gupta, H., Vahid Dastjerdi, A., Ghosh, S.K., Buyya, R.: ifogsim: A toolkit for modeling and simulation of resource management techniques in the internet of things, edge and fog computing environments. Softw. Pract. Exp. **47**(9), 1275–1296 (2017)
9. Iturriaga, S., Nesmachnow, S., Bernabe, D., Bouvry, P.: Energy efficient scheduling in heterogeneous systems with a parallel multiobjective local search. Comput. Inform. **32**(2), 273–294 (2013)
10. Kocot, B., Czarnul, P., Proficz, J.: Energy-aware scheduling for high-performance computing systems: a survey. Energies **16**(2), 890 (2023)
11. Krzhizhanovskaya, V.V., Korkhov, V.V.: Dynamic load balancing of *black-Box* applications with a resource selection mechanism on heterogeneous resources of the grid. In: Malyshkin, V. (ed.) PaCT 2007. LNCS, vol. 4671, pp. 245–260. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-73940-1_26
12. Kwok, Y.K., Ahmad, I.: Static scheduling algorithms for allocating directed task graphs to multiprocessors. ACM Computi. Surv. (CSUR) **31**(4), 406–471 (1999)
13. Luo, P., Lü, K., Shi, Z.: A revisit of fast greedy heuristics for mapping a class of independent tasks onto heterogeneous computing systems. J. Parallel Distrib. Comput. **67**(6), 695–714 (2007)
14. Mach, P., Becvar, Z.: Mobile edge computing: a survey on architecture and computation offloading. IEEE Commun. Surv. Tutor. **19**(3), 1628–1656 (2017)

15. Maheswaran, M., Ali, S., Siegel, H.J., Hensgen, D., Freund, R.F.: Dynamic mapping of a class of independent tasks onto heterogeneous computing systems. J. Parallel Distrib. Comput. **59**(2), 107–131 (1999)
16. Mao, Y., You, C., Zhang, J., Huang, K., Letaief, K.B.: A survey on mobile edge computing: the communication perspective. IEEE Commun. Surv. Tutor. **19**(4), 2322–2358 (2017)
17. Nesmachnow, S., Dorronsoro, B., Pecero, J.E., Bouvry, P.: Energy-aware scheduling on multicore heterogeneous grid computing systems. J. Grid Comput. **11**(4), 653–680 (2013)
18. Sajid, M., Raza, Z., Shahid, M.: Energy-efficient scheduling algorithms for batch-of-tasks (bot) applications on heterogeneous computing systems. Concurr. Comput.: Pract. Exp. **28**(9), 2644–2669 (2016)
19. Shi, W., Cao, J., Zhang, Q., Li, Y., Xu, L.: Edge computing: vision and challenges. IEEE Internet Things J. **3**(5), 637–646 (2016)
20. Sonmez, C., Ozgovde, A., Ersoy, C.: Fuzzy workload orchestration for edge computing. IEEE Trans. Netw. Serv. Manag. **16**(2), 769–782 (2019)
21. Tarafdar, A., Debnath, M., Khatua, S., Das, R.K.: Energy and makespan aware scheduling of deadline sensitive tasks in the cloud environment. J. Grid Comput. **19**(2), 19 (2021)
22. Topcuoglu, H., Hariri, S., Wu, M.Y.: Performance-effective and low-complexity task scheduling for heterogeneous computing. IEEE Trans. Parallel Distrib. Syst. **13**(3), 260–274 (2002)
23. Zhang, Q., Li, H.: MOEA/D: a multiobjective evolutionary algorithm based on decomposition. IEEE Trans. Evol. Comput. **11**(6), 712–731 (2007)
24. Zhdanov, D., Korkhov, V.: On algorithms of task placement in computing clusters taking into account task and system topologies. Phys. Part. Nucl. **55**(3), 624–626 (2024)
25. Zhou, A., Qu, B.Y., Li, H., Zhao, S.Z., Suganthan, P.N., Zhang, Q.: Multiobjective evolutionary algorithms: a survey of the state of the art. Swarm Evol. Comput. **1**(1), 32–49 (2011)
26. Zitzler, E., Laumanns, M., Thiele, L.: Spea2: improving the strength pareto evolutionary algorithm. TIK Rep. **103** (2001)

Creating an Operating System for Video Camera Chips in an Intelligent Video Surveillance System

Alexey Subbotin[1](✉), Nataly Zhukova[2], and Elena Stankova[3]

[1] Saint-Petersburg State Electrotechnical University, Professor Popov Street 5, 197376 St. Petersburg, Russia
alesu1543@gmail.com

[2] St. Petersburg Institute of Informatics and Automation of the Federal Research Center at the Russian Academy of Sciences (FRC RAS), 14th Line V.O. 39, 199178 St. Petersburg, Russia

[3] St. Petersburg State University, 7-9, Universitetskaya Nab., 199034 St. Petersburg, Russia

Abstract. This article discusses a method for creating an operating system for devices with limited performance, specifically video camera chips used in intelligent video surveillance systems. Analogs are considered: general-purpose operating systems and special-purpose operating systems for devices with limited computing resources. The problem of fast launching of an operating system on process computers in devices is identified. Applications in an operating system can run faster if the operating system itself is designed for a single task, taking into account the architectural features of the embedded processor in the chips of a technical device, such as a video camera. Deep integration of the control program, file system, and scripting language in the operating system kernel is possible. The kernel of such a system runs faster for intelligent video surveillance systems. A solution is proposed, involving development of a custom operating system for video camera chips with limited computing resources. The new operating system runs and launches applications 67% faster in real time mode and boots twice faster than similar systems. The proposed method for increasing the speed of program execution in the operating system environment and for quickly launching the new operating system on video surveillance devices can also be applied in other devices, such as routers, tablets, and embedded computers for a wide range of applications.

Keywords: Operating systems · Technological computers · Video camera chips · Intelligent video surveillance systems

1 Introduction

An operating system is essential for any computer. Starting with the Multics operating system (https://www.multicians.org/), all operating systems provide a basic set of input/output operations and are an extension of the BIOS (Basic Input/Output System). A BIOS operates only with the machine, while an operating system is focused on a

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 511–523, 2027.
https://doi.org/10.1007/978-3-032-30530-5_32

specific computer architecture, particularly the central processing unit (CPU). Embedded operating systems (OS) for devices are quite common and not something specific. Most embedded OS do not have user-facing capabilities at the kernel level; they are resource managers and task schedulers. New devices are added to the operating system's address space. In addition to working with devices, the operating system divides RAM for shared use between applications. Each operating system contains a basic set of utilities and commands for working with devices and applications. Interrupts can be both hardware and software.

The main task of an operating system is to delimit memory (stack in CPU, heap in RAM, disk in hard drive) and assign access rights. It can be said that the operating system is a RAM manager and utilities [1]. Each OS executes application code with additional information contained in the header. Application headers can be of various formats: PAGEZERO for MacOS, MZ for Microsoft Windows, ELF for Linux. The bytecode of different applications runs concurrently in the OS and should not interfere with each other. It is possible to read the RAM of all applications, but in some cases (like *BSD), reading the memory of other applications is difficult or even impossible (with hardware memory encryption).

1.1 History of Operating Systems

The original function of all operating systems was to redirect input/output, while simultaneously performing arbitrary operations on the data. The concept of redirecting data streams from one device (the transmitter) to another (the receiver) has been preserved in all computers and operating systems (Fig. 1). The Multics operating system, with its graphical interface (a screen divided into four sections with a menu at the top), was designed for US missile defense. Multics also worked for AT&T, which provided telephone and telegraph services to the US population. Conversations were recorded on reels, and Multics commands helped US officers to extract necessary conversation fragments (files) into a spooler (a separate storage reel with audio interfaces on a GE-645 computer running Multics). The principle of file transfer and data redirection was inherited by the UNIX operating system, which source code is ten times less than the source code of Multics.

Despite the importance of a computer (or mainframe) operating system, creating your own isn't all that difficult. The popular Minix OS is a UNIX clone, the volume of the source code of which is about a half of a book volume (1,200 pages). However, Minix is full-fledged operating system. Linux was created as a Minix clone with added networking features.

1.2 Chips in Video Cameras

It's difficult to impress anyone with grain-sized chips these days, since everyone uses bank cards, and the chip on the card is a fully-fledged computer. However, a computer inside a device (for example, a video camera) can significantly outperform bank chips. It may have relatively large RAM and ROM, a good processor with integrated graphics functions [20, 21]. Video can be processed not only in the cloud or on a fog video

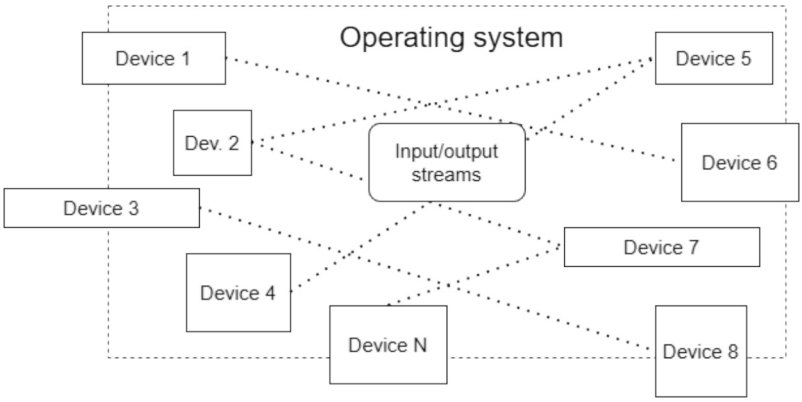


Fig. 1. The principle of input-output redirection in the OS.

server, but also directly on the camera's process computer using an integrated system-on-board (SoC) processor. ARTPEC processors for Axis Communications come in various series and modifications (chip generations), which differ in capabilities and performance (depending on the functions they can perform for video cameras). Built-in functions include color correction, haze removal, night mode, clarity, brightness and contrast adjustment, and many other features supported in explosion-proof and waterproof Axis video cameras.

2 Overview of General-Purpose Operating Systems

A popular operating system for embedded industrial computers is Linux (https://kernel.org/), which is a monolithic kernel with device drivers, utilities, and file systems. Various Linux-based OS builds are available, including Ubuntu, Astra, ALT, SUSE, and many others, differing primarily in their application sets. Linux supports basic UNIX commands (including those inherited from Multics), allowing it to run the simple X-Window windowing system (https://x.org/) and the more complex KDE (https://kde.org/) or Gnome (https://www.gnome.org/) on top, which contain a large number of additional applications and are more demanding on the graphics adapter (GPU). In the absence of a powerful graphics adapter, for example, in an industrial PC, it is recommended to use the basic X-Window windowing system or the command line. Linux supports various processor (CPU) architectures, a feature that no other OS can provide. At the time of writing, Linux supports the following processors: ARM, x86, PowerPC, RiscV, MIPS, and SuperH [18].

The most popular application operating system is Microsoft Windows (https://support.microsoft.com/ru-ru/windows), thanks to its simplicity and accessibility (it's affordable, and in some cases, Microsoft Windows comes pre-installed almost free). Windows has a robust API (Application Programming Interface) with high-quality documentation for developing applications. Windows supports threaded processes with priorities, allowing multiple programs from the steak and heap to run simultaneously. Priorities allow

for a real-time effect, where a program effectively takes over all processor resources and provides almost instantaneous response times, which is commonly considered real-time.

The macOS family from Apple (https://www.apple.com/os/macos/) accounts for a large amount of desktop and smartphone operating systems. Mac OS versions for PCs and iPhones were not quite stable, but after the FreeBSD kernel was implemented with permission from the developers, the macOS family of operating systems became extremely reliable. This OS line is very popular in Apple devices (smartphones, laptops, smartwatches, tablets, game consoles, professional video cameras, and televisions) due to Apple's licensing policy, which almost never sells macOS to non-Apple device makers [3, 4].

The *BSD OS line from the University of California, Berkeley (https://www.freebsd.org/) is considered the most stable OS. Various modifications exist: FreeBSD with support of Linux applications; NetBSD for routers; OpenBSD with strict licensing requirements for application development. Memory encryption makes it virtually impossible to modify other programs. A FreeBSD application runs only in its own memory, allocated by the OS. If the application crashes, the FreeBSD kernel restarts the application. This heap memory policy creates significant challenges for application developers, but the application will operate reliably and with high stability [15–17].

There are operating systems that are not widely used (Chrome OS, ReactOS, Haiku, OpenSolaris), but are designed for specific devices in the interests of individual corporations. The UNIX-like operating system OpenSolaris v.134 (http://www.opensolaris.com/) from Oracle has its own kernel and graphical interface (MilaX, marTux, Nexenta, etc.) developed for SQL and NoSQL data servers. OpenSolaris has its own DBMS, OracleDB (https://www.oracle.com/database/). OpenSolaris is also used on backend servers with programming languages: Java, Perl, PHP from the Web Stack package with MySQL, and Apache. Google's Chrome OS v.140.0.7339.201 (https://chromeos.google/) is used for tablets running the Chromium-based Google Chrome browser. Chrome OS supports ARM or x86 processors. Chrome OS is primarily installed on tablets such as the Google Pixel Slate, Asus Chromebook Tablet CT100, Acer Chromebook Tab 10, HP Chromebook X2, and Samsung Galaxy Chromebook 2.

Comparison of operating system capabilities is presented in Table 1. The systems are compared based on the following capabilities: p1 – operation under different processor types; p2 – reliable operation of the RAM manager (heap); p3 – applicability (or version availability) for technological embedded computers; p4 – reliable separation of user rights; p5 – configuration and editing of the kernel source code; p6 – implementation of process threads for the real processor operation mode (the response time of the first process is less than 4 ms); p7 – the ability to quickly start the OS without checking devices and loading drivers for each device in the PC; p8 – compilation and execution of binary code of applications from other OS; p9 – support for UNIX, Multics commands and the POSIX standard in general for executing the X-Window program; p10 – support for command line utilities inherited from Multics and UNIX; p11 – the ability to disable the multi-user mode of the OS and use the auto login function to speed up loading the desktop (graphical interface in the OS).

Table 1. Comparison of the features and limitations of the considered operating systems.

Name of the system	p1	p2	p3	p4	p5	p6	p7	p8	p9	p10	p11
Linux	+	-	+	+	+	+	-	+	+	+	+
Windows	-	-	+	+	-	+	+	-	-	-	+
macOS	-	-	+	+	-	-	+	+	-	+	+
FreeBSD	-	+	+	+	+	-	-	+	+	+	+
ReactOS	-	-	-	-	-	-	-	-	n/a	-	n/a
Haiku	-	-	-	-	-	-	-	-	n/a	-	-
OpenSolaris	-	-	-	-	-	-	-	-	+	+	+
QNX	-	+	+	+	-	+	-	+	+	+	+

According to the results of analyzes of existing systems (Linux, Windows, macOS, FreeBSD, ReactOS, Haiku, OpenSolaris, QNX) there is no system that supports all the required features.

3 Overview of Operating Systems for Devices

Embedded Linux (https://ubuntu.com/embedded) is a very popular operating system for chips, embedded devices, and industrial computers. It does not have the X-Window graphical shell and powerful KDE or Gnome window systems [19]. This version of Linux contains only three components: Bootloader (boot tools), Kernel (memory manager for applications, device drivers), and RootFS (a set of file systems and network utilities). Embedded Linux compilation and build tools (Toolchain, Yocto, Buildroot) can be installed separately [6].

QNX (http://qnx.com/) is designed for embedded systems and operates in real-time. This means that the first application launched in QNX responds within milliseconds, while other applications experience significant lag. QNX is used in train control systems and other systems that require guaranteed, near-real-time response. According to official documentation, real-time mode in QNX 6.3 is defined as a guaranteed application response time of up to 4 ms when using a processor with a frequency of 40 MHz or higher [10, 11].

Windows IoT or Windows Embedded OS (https://developer.microsoft.com/en-us/windows/iot/) is for embedded technological (specialized) devices: cash registers, self-service kiosks, ATMs, payment systems, and other technological and industrial devices. For example, for the ICOP Vortex 86DX computer, which is a SoC (system on a chip) [13].

NetBSD (https://netbsd.org/) has proven itself as an OS for routers and network devices. It is highly reliable and built on the BSD kernel, excluding large software modules that are unnecessary for technological computers based on chips for industrial devices [12]. MacOS (https://support.apple.com/en-us/macos) is used in IoT devices certified by Apple [14].

The presented operating systems are compared in Table 2. The operating systems are compared based on the following features: p1 – support of mobile devices and smartphones; p2 – support of Internet of Things technology; p3 – support of industrial devices and production control; p4 – support of X-Window for running applications in the KDE/Gnome graphical environment or browser; p5 – kernel customization for use in devices with limited CPU and GPU; p6 – applicability for intelligent video surveillance system video camera chips; p7 – availability of implemented protocols for working with video servers.

Table 2. Comparison of the features and limitations of the considered systems.

Name of the system	p1	p2	p3	p4	p5	p6	p7
Embedded Linux	+	n/a	+	+	+	+	+
QNX	+	-	+	+	-	-	-
Windows IoT	+	+	-	+	-	+	+
NetBSD	-	-	+	-	-	-	-
macOS	+	+	-	-	-	n/a	+

According to the results of analyzes of existing systems (Embedded Linux, QNX, Windows IoT, NetBSD, macOS) there is no system that supports all the required features.

4 Operating System Problems

After loading the file system driver, the OS checks the PC's devices. A timeout (device response time) is specified for each device type in the OS configuration files. The OS sends an initialization signal each time, obtains the series and model of the device (video adapter, network card, card reader, etc.), and checks for drivers (from the heap if it's Linux; from disk if it's QNX or Windows). If the required driver is not found, it displays an unknown device message [7]. After launching the application in some cases bytecode translation is required for a different processor (for example, Linux kernel was originally designed for the x86 processor architecture, so is requires bytecode translation) (Fig. 2).

It's possible to create a specialized OS for video camera chips that will run applications faster and boot up faster. The slow boot problem can be solved using Puppy Linux 9.5 (http://www.puppylinux.com/), which bypasses device checks and immediately opens the desktop with a window system. If a device fails, the OS user will find it out during the OS's operation.

Older motherboards and devices built before 2001 lacked the ability to continuously supply power. These devices used AT power supplies, not ATX (Advanced Technology Extended). The motherboard and the entire computer or device (camcorder or embedded computer) were not powered after a power outage. After powering on, each device

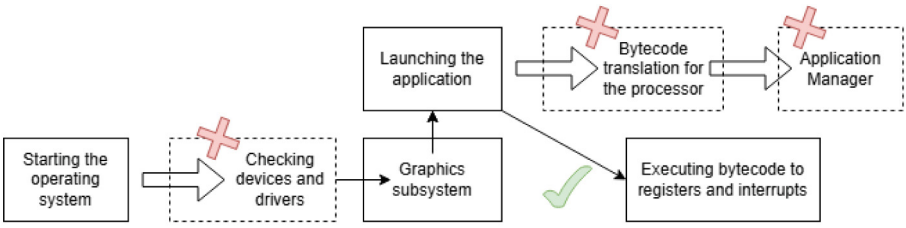


Fig. 2. Visual representation of OS problems.

received a [READY] command from the OS, which took time. Then hibernation mode was introduced, eliminating the need to poll devices. Now, there's no need to put the computer's devices into a ready state, since they were already polled when the power was turned on. If a device fails, a message will be displayed for the operating time of the computer or camcorder.

The problem of slow program execution in the OS can be solved by not converting applications in RAM, as in Linux, but executing them directly, as in FreeBSD. Therefore, binary emulation of Linux programs (ELFs) runs faster in FreeBSD than on the native 7.0.3 kernel (rsync://rsync.kernel.org/pub/), which both support the POSIX (Portable Operating System Interface) standard, which consists of basic utilities and commands. Many commands are inherited by the IEEE POSIX Certification standard (https://posix.opengroup.org/) from Multics and UNIX. In Linux, the commands are contained in the kernel itself with drivers. For example, in version 3.2.0, at startup, the 4.7 MB file "vmlinuz-3.2.0–53-generic" is read from the "/boot/" directory and written to a virtual disk in RAM. The directory also contains additional utilities with the prefixes: "abi-Ver.", "config-Ver.", "initrd.img-Ver.", "System.map-Ver." A 14 MB version 3.2.0 file, but the Linux kernel is still considered monolithic. The QNX4 operating system (https://blackberry.qnx.com/) contains a microkernel only 4 KB in size. The remaining OS functions in files: Fsys, Proc, Dev, Net, IPC, interrupt handling, and others are loaded by the microkernel when needed. There is no need to recompile the kernel, like in Linux. Even the superuser "root" cannot enter the "boot" directory and delete data, only "boot." QNX 6 Neutrino (https://qnx.software/) is a single-process (node) philosophy that blocks all other processes for itself (up to 4 ms), and other processes significantly slow down. Windows 98 operated on this principle, as it didn't have processor threads like Windows NT/XP. However, a single Windows 98 process didn't guarantee a response time of up to 4 ms, unlike QNX4. Starting with Windows NT, high-priority emulation and real-time mode are supported. Simply launch any program with the "start /high prog" or "start /realtime prog" commands. In FreeBSD, the kernel automatically compiles to an ELF file and runs. Possible make settings in the "/usr/src/sys/conf/kern.*.mk" config files during the initial FreeBSD boot from the "/red/herring" file using the "vi" and "ee" editors. The ELF file will then be compiled and launched. The kernel then loads additional services and even a terminal, if needed. Kernel parameters are stored in the "CONFIG_FILE" file. The kernel is launched with the "sh" command. Using heap fragmentation for a single task will speed up application execution, but this requires developing a custom OS.

If the OS is designed for a specific task [2, 5] (intelligent video surveillance for video camera chips), then such an OS will boot and run faster.

5 Solution Method

Despite its complexity (and its importance for PCs), creating an operating system is simple. The volume of the source code of the popular Minix 3.3.0 OS (https://minix3.org/) is merely as the volume of appendix in Andrew Tanenbaum's book "Creating an Operating System" [1]. Samsung monitors contain their own Tizen OS (in earlier versions: Bada and Orsay OS) with regular updates. Not all new OS are Linux clones, like Android Linux. Creating your own OS kernel is not that difficult.

A full-fledged OS can be considered as a program that provides:

1) reading the binary application format;
2) RAM manager for application collaboration;
3) memory management (heap, steak, disks) with access rights;
4) support of various devices with data stream redirection (event or interrupt pool for applications).

The kernel can be monolithic (as in Linux), where device drivers, memory manager, and utilities are contained in a single file and transferred to the RAM (heap) at OS startup. The kernel can also be a microkernel (as in QNX) and contain multiple files that are loaded separately.

The quality of an embedded technological operating system for video surveillance devices can be improved based on the following [9]:

- scan devices during OS operation, not at boot time;
- develop a RAM manager to allocate the heap to a single application;
- execute bytecode from an application without translation to a specific processor architecture;
- optimize bytecode for RAM clustering (heap) and the registers of a specialized processor in the device chip (video camera);
- eliminate permissions, access restrictions, and process threads.

To interact with the operating system, a command-line terminal capable of entering commands via familiar protocols is required, as connecting directly to the camera's pins and interfaces on the board is very inconvenient. This is especially true if the camera is located remotely (on the ceiling, at a height, etc.). A management tool and application development for remote connection are needed. The browser-based shell should be deeply integrated, such as Webmin 2.630 (https://webmin.com/). The tool should be deeply integrated into Subix with the ability to connect remotely via TCP/IP and combine the traditional Webmin push-button interface with the Command Line Interface (CLI) with four components (Prompt, Console, Terminal, Shell) in a simple and straightforward manner for a camera configuration specialist, not a system developer. This ensures easy daily use and configuration as needed for intelligent video surveillance systems.

In the WinSubs [8] graphics subsystem of the Subix operating system it is sufficient to create a simulated console with execution of all commands and viewing video fragments

from the camera's flash memory. A possible implementation option for the graphical subsystem (Fig. 3) for the Subix operating system is an area in the browser where the data is displayed and a graphical shell with a command line developed in the form of a website using HTML, CSS 3.2, and JavaScript as the frontend and PHP with Subscript [2] as the backend. The workspace is divided into two parts: in the upper part the Subix user is provided with real-time video viewing (left) and viewing individual fragments from past periods (right) with the ability to navigate using arrows that change the chronology depth (center). The second half (bottom) contains a console in which commands can be entered even without the WinSubs graphical subsystem (left). On the right is the area of control buttons: Menu for returning to the main website screen; Settings for switching to a separate video camera settings screen. The cross at the top of the Subix logo opens a page for rebooting the system.

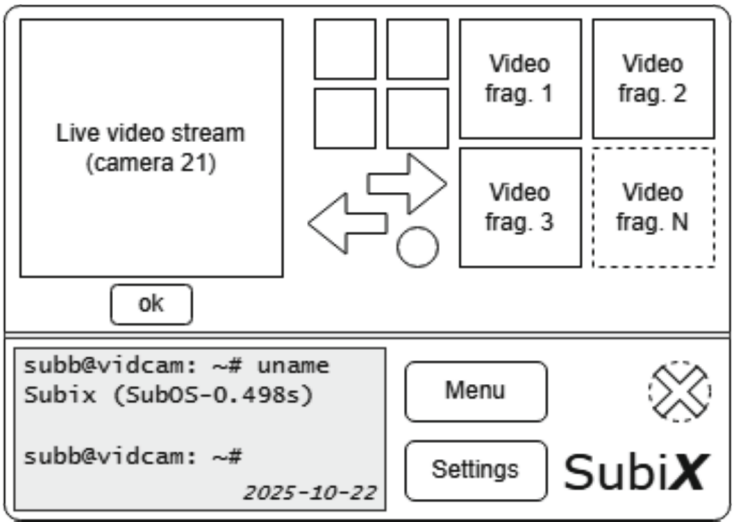


Fig. 3. Graphic subsystem for Subix OS.

In addition to interrupt remapping, executing instructions from the heap, maintaining access rights, drivers, and working with devices, any OS needs to boot from disk, which requires a file system (FS). For Axis P1365 MK II video cameras on chips with the ETRAX 100LX processor with RISC ISA instructions and 16 MB of RAM, the SubSysVideo FS [7] is suitable. It is adapted for intelligent video surveillance systems and is almost 50% faster than similar systems. WinSubs [8] is a suitable graphical shell; it visualizes events and is 20% faster than similar systems. For creating applications in the Subix operating system (SubOS-0.498s), the SubScript [2] scripting language is suitable; it is 6 times faster than similar systems. A description and comparison of the SubScript language [2] with a description of the language execution model (JIT, AOT, interpreter), memory management and comparative performance tests with Lua, Python, PHP and more than 17 languages (https://github.com/alex1543/17langs) on the same hardware are presented in detail in the work [2]. A new scripting language with

certain characteristics (the absence of configuration files, plug-in libraries and a complex, intricate scripting language core) and the use of the GCC 15.2 compiler (https://gcc.gnu.org/) can achieve an increase in the data exchange speed. The SubScript language [2] operates on the principle of macros and templates, for example, like rss2fido Script (https://rss2fido.sf.net/). The SubScript scripting language is created specifically for storing data for a single task and is executed as efficiently as possible. Performance measurements were carried out through JavaScript on the website (https://github.com/alex1543/EditTable), where you can select different languages for comparison (PHP, Perl, Python, Ruby, Go, Lua, Elixir, Dart, Node.js, etc.) for working on the backend for visualizing events and storing data in an intelligent video surveillance system [3, 4, 8].

6 Case Study

The authors of the study developed a remote connection application for the Subix OS (Fig. 4), which was used at the State Unitary Enterprise "Petersburg Subway" (https://metro.spb.ru/). Subix OS runs a webserver written in C for the GCC compiler (https://github.com/alex1543/practGCC). To connect to the graphical interface, simply enter "https://192.168.25.147/subix" in your browser, where "192.168.25.147" is the IP address, which can be changed in the settings. Connection is possible via either twisted pair (Ethernet) or USB.

After conducting the experiment (Table 3) within which the following parameters were measured: p1 - OS boot time; p2 - program execution time from RAM (heap); p3 - OS interrupt response time, it was found that Subix OS (SubOS-0.498s) works 2.3 (2.258) times faster than its analogs (Embedded Linux, QNX, Windows IoT, NetBSD, macOS). Data on the p2 and p3 parameters for QNX and Subix (SubOS-0.498s) could not be measured more accurately due to the timer limitation to 1000 ms. Launching and running applications in the new Subix operating system are 67% faster than QNX OS in real time. On other similar OSs, it was not possible to achieve real time mode, therefore the p2 and p3 parameters show the result of OS operation, that is on average slower by 978 (977.6) ms. For one application execution and launch.

To reproduce the experiment, compile the Subix OS source code (SubOS-0.498s) using a BASH script ("./comp_subb.sh"). Install all dependent applications, measure the time per day (24 or 25 h), open the logs in Apache OpenOffice (https://www.openoffice.org/), run JavaScript, and fill the resulting table. Thirty-one Axis P1365 IP video cameras (https://help.axis.com/en-us/axis-p1365-mk-ii) were used for testing. Statistics were processed on an Apple iMac M4 computer with the following technical specifications: 4 cores, IPS, 4.5K Ultra HD (4480x2520), 24 GB LPDDR5X, 1000 GB SSD, macOS 26.1 (T2/Apple Silicon 2025).

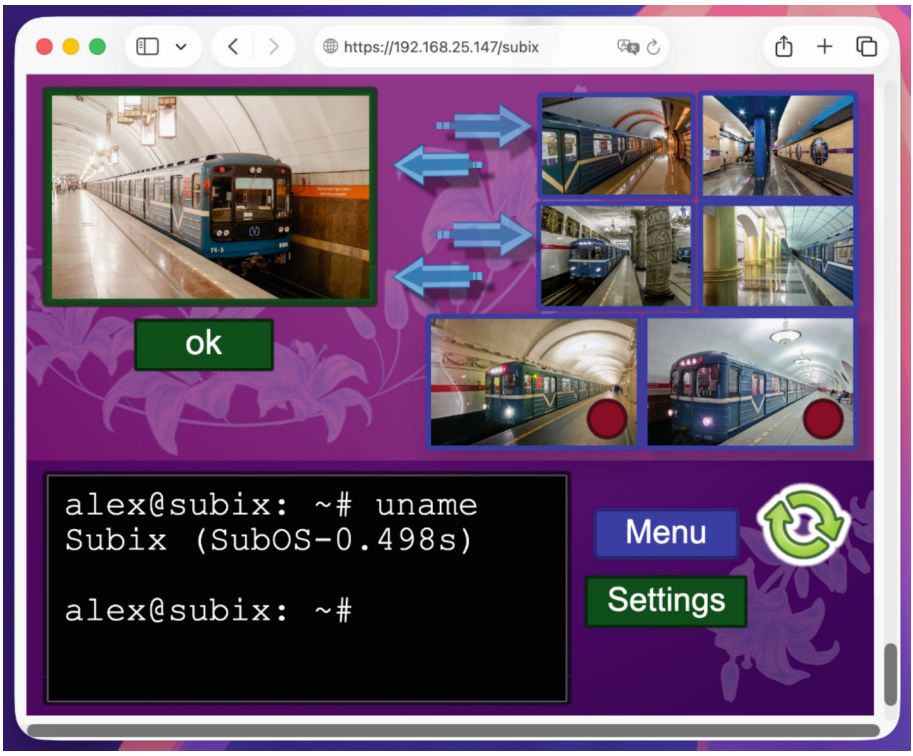


Fig. 4. Implementation of the graphical subsystem and command console in the Subix OS.

Table 3. Correlation of OS names and performance parameters.

OS Name	p1 (sec.)	p2 (ms.)	p3 (ms.)	Slower than Subix (times)
Embedded Linux[a]	157	2174	1279	2.63
QNX	124	2	1	2.03
Windows IoT	139	1492	873	2.32
NetBSD	142	1759	1028	2.37
macOS	117	893	275	1.94
Subix (SubOS-0.498s)	61	1	1	-
Subix is faster on average:				2.258

a) The default system in the subway (now).

7 Conclusion

The authors have created an operating system adapted for video camera chips for intelligent video surveillance systems.

The study found that the new Subix OS (SubOS-0.498s) boots twice as fast (2.22623) as its competitors; applications launch from disk 1000 times faster (1264) and execute in RAM 600 times faster (691.2). In real-time mode, the new operating system runs and launches applications 67% faster and in non-real-time is 600 times faster. According to the QNX documentation in real-time mode the first launched application must respond within 4 ms for a processor of 40 MHz or faster.

The proposed method for improving OS performance for embedded process computers can be applied not only in video cameras of video surveillance systems, but also in other devices: smart routers, industrial controllers, and process tablets at various stages of production and technical system maintenance.

Acknowledgments. This work was supported by the state budget, project No. FFZF-2025–0019.

References

1. Tanenbaum, A.S., Woodhull, A.S.: Operating systems: design and implementation, January 2006. ISBN 9780131429383
2. Subbotin, A., Delhibabu, R., Zhukova, N.: A script language creation for event visualization and data storage in an intelligent video surveillance system. Adv. Artif. Intell. Mach. Learn. **4**(4), 3161–3172 (2024)
3. Osipov, V., Zhukova, N., Subbotin, A., et al.: Intelligent escalator passenger safety management. Sci. Rep. **12**, 5506 (2022). https://doi.org/10.1038/s41598-022-09498-x
4. Tianxing, M., Vodyaho, A., Zhukova, N., Subbotin, A., et al.: Urban intelligent assistant on the example of the escalator passenger safety management at the subway stations. Sci. Rep. **13**, 15914 (2023). https://doi.org/10.1038/s41598-023-42535-x
5. Osipov, V., Nikiforov, V. Recurrent neural networks with controlled elements in restoring frame flows. Inf. Control Syst., 10–17 (2019). https://doi.org/10.31799/1684-8853-2019-5-10-17
6. Simmonds, C.: Mastering Embedded Linux Programming (2017)
7. Subbotin, A.N., Zhukova, N.A.: Development of a file system for storing data of an intelligent video surveillance system. Sci. Tech. J. Inf. Technol. Mech. Opt. **25**(2), 253–260 (2025). (in Russian). https://doi.org/10.17586/2226-1494-2025-25-2-253-260
8. Subbotin, A., Zhukova, N., Stankova, E.: Intelligent browser window system for visualizing events detected from video in distributed fog computing environment. In: Gervasi, O., et al. Computational Science and Its Applications – ICCSA 2025 Workshops. ICCSA 2025. Lecture Notes in Computer Science, vol. 15894, pp. 309–324. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97648-3_21
9. Titova, A.P., Karpov, N.S., Alexandrov, I.A.: Creation of an operating system that enables additive manufacturing. In: Proceedings of the 2024 Dynamics of Systems, Mechanisms and Machines (Dynamics), Omsk, Russian Federation, pp. 1–5 (2024). https://doi.org/10.1109/Dynamics64718.2024.10838674
10. Ilyushin, P., Volnyi, V., Suslov, K.: Review of approaches to creating control systems for solid-state transformers in hybrid distribution networks. Appl. Sci. **15**(20), 10970 (2025). https://doi.org/10.3390/app152010970
11. Berry, F., Kleihorst, R., Galàn, R.C.: Smart camera architecture. J. Syst. Architect. **59**(10), Part A, 817 (2013). ISSN 1383-7621, https://doi.org/10.1016/j.sysarc.2013.11.002

12. Gao, L., Zong, R.: Design of intelligent window automatic monitoring system based on microcontroller control. J. Electron. Res. Appl. **8**, 34–40 (2024). https://doi.org/10.26689/jera.v8i4.7922
13. Said, Y., Saidani, T., Atri, M.: Embedded architecture for real-time video processing. Imaging Rob. **15**(4) (2015). ISSN 2231-525X, https://doi.org/10.1109/DSD.2008.156
14. Lee, J.-H., Hyeon, B.-S., Jeon, O.-Y., Park, N.I.: Analysis of real-time operating systems' file systems: built-in cameras from vehicles. Forensic Sci. Int. Digit. Invest. **44**(Supplement), 301500 (2023). ISSN 2666-2817, https://doi.org/10.1016/j.fsidi.2023.301500
15. Zhou, Q.: Frequency-domain characterization of finite sample linear systems with uniform window inputs. Signals **6**, 1 (2025). https://doi.org/10.3390/signals6010001
16. Pietzsch, A., Woerd, J., Andrae, M., Gebbeken, N.: Analysis of the load-bearing behavior of catcher-cable systems for hazard prevention for windows under blast loads (2024)
17. Elavarasi, Anand, M., Sathya, S.: Design and implementation of reconfigurable virtual stage for real-time component using FreeRTOS. Int. J. Pure Appl. Math. **117**(16), 719–726 (2017). ISSN: 1314-3395, https://doi.org/10.12733/jica.2017.117.16
18. Sisi, C.: Empirical study on functional defects in multi-window Android systems. Appl. Comput. Eng. **73**, 288–294 (2024). https://doi.org/10.54254/2755-2721/73/20240413
19. Giatsintov, A., Mamrosenko, K., Bazhenov, P.: Architecture of the graphics system for embedded real-time operating systems. Tsinghua Sci. Technol. **28**(3), 541–551 (2023). https://doi.org/10.26599/TST.2022.9010028
20. Deng, L., Shi, Z., Wang, Y., Yu, X., Guan, Y., Xu, Z.: Hybrid window decoding for joint source channel anytime coding system. Entropy **26**, 940 (2024). https://doi.org/10.3390/e26110940
21. Wolf, W., Ozer, B., Lv, T.: Smart cameras as embedded systems. Computer **35**(9), 48–53 (2002). https://doi.org/10.1109/MC.2002.1033027

Sustainable evolution of long-Distance frEight and paSsenger Transport (SOLIDEST 2026)

Comparison of Container Routing in Liner Shipping Networks

Massimo Di Gangi, Orlando M. Belcore(✉), and Antonio Polimeni

Department of Engineering, University of Messina, Messina, Italy
{mdigangi,obelcore,antonio.polimeni1}@unime.it

Abstract. The increasing volatility of liner shipping networks, characterized by frequent capacity reallocations and schedule disruptions, calls for robust and comparable container routing models. This paper presents a systematic comparison of two approaches: the Connection Scan Algorithm (CSA) and the RAPTOR algorithm. To enable a model-independent assessment, routes are projected onto a canonical representation based on sequences of ports and liner services. Using a set of scenarios to increase network complexity, routing outcomes, computational performance, and route similarity between the two approaches are evaluated. The results show that CSA and RAPTOR exhibit greater sensitivity to network dynamics, generating structurally different routes with more transshipments, particularly in large and volatile scenarios. The proposed framework provides a unified basis for comparing heterogeneous routing models and highlights trade-offs between computational efficiency, temporal realism, and robustness, offering guidance for both strategic planning and operational analysis in modern liner shipping networks.

Keywords: container shipping · liner shipping · connection scan algorithm · RAPTOR · maritime transport

1 Introduction

Liner shipping networks constitute the backbone of global maritime transport, enabling the movement of containers across intercontinental routes [1, 2]. The planning and operational management of these networks require a detailed understanding of port management [3, 4], container routing [5, 6], transshipment processes [7, 8], and schedule synchronization [9, 10]. Container routing decisions directly affect transit times, reliability, and the utilization of maritime and terminal resources. Therefore, container routing represents a critical decision problem in maritime logistics, where containers are assigned to sequences of liner services and ports subject to service schedules, transshipment constraints, and operational requirements.

Recent developments in liner shipping operations further underscore the need for robust and temporally consistent routing models. According to recent analyses by Sea Intelligence [11], 2025 represents a turning point in ocean liner policy, distinguished by substantial changes in vessel deployments across major East–West trade lanes. This volatility is captured by the notion of operational capacity churn, defined as the sum of

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 527–542, 2027.
https://doi.org/10.1007/978-3-032-30530-5_33

vessel capacity entering and exiting a trade route over time, which serves as a proxy for structural instability in service deployment. Such dynamics are different from the prepandemic paradigm, in which container flows followed relatively predictable patterns. Instead, liner shipping networks increasingly resemble dynamic systems subject to short-term demand fluctuations, operational disruptions, and rapid reconfiguration of services. Routing decisions based on static or time-aggregated network representations risk overlooking critical feasibility and synchronization constraints. These developments raise fundamental questions regarding the adequacy of different network modeling approaches for simulating/designing container routing. An assessment of the adequacy of line-based representations under operational volatility is required to understand whether time-expanded formulations are needed (and more generally, time-sensitive routing approaches) to identify feasible itineraries, transfer patterns, and rerouting opportunities.

In this context, this study provides a comparison of alternative models for container routing. By designing routing solutions onto a common canonical representation, the study quantifies differences in routing outcomes, route structure, thereby offering insights into the trade-offs between modeling simplicity and operational realism in liner shipping networks.

Routing in liner shipping can be represented at different levels of detail, ranging from service-oriented models to fully time-dependent formulations that explicitly incorporate port calls and vessel schedules. Therefore, selecting an appropriate modeling approach is crucial for both strategic planning and operational decision-making, as it affects the realism of routing outcomes and the computational tractability of the problem. Network representations that explicitly account for time (such as time-expanded networks) or connection-based approaches are particularly well suited to this purpose, as they enable the identification of alternatives and the assessment of their impact on transit times and transshipments. In addition to connection-based approaches, round-based timetable routing algorithms (e.g., RAPTOR [12]) could represent a natural extension of the methodological framework considered in this study. Although a large body of literature addresses container routing and transshipment optimization [13, 14], systematic comparisons of routing approaches remain scarce, particularly with respect to:

- differences in routing decisions and transshipment patterns,
- structural similarities and divergences of the resulting routes,
- computational scalability across increasing network complexity,
- robustness under operational disruptions and re-routing scenarios.

To try to fill this gap, this research tries to answer the following questions:

- How do different temporal routing algorithms (CSA and RAPTOR) perform under increasing network complexity in liner shipping?
- What are the trade-offs between these routing approaches in terms of route quality, computational feasibility, and robustness?
- Does a unified, model-independent framework facilitate a systematic comparison and highlight key differences relevant for strategic and operational decision-making?

Routes generated using connection-based approaches and RAPTOR algorithms are projected onto a common canonical representation, enabling direct and quantitative comparison. Similarity metrics, performance indicators, and robustness analysis are

employed across a range of network scenarios, providing insight into the trade-offs between modeling goodness and computational efficiency for both strategic and operational applications. The findings are expected to be of interest for both researchers in advancing routing methods and practitioners aiming to optimize network design, enhance robustness to disruptions, and improve efficiency in container transport operations. Thus, the paper serves both as a methodological framework and as a practical guide for stakeholders.

The paper is structured as follows. Section 2 reports a brief literature review on modeling and approaches for route simulation and design. Section 3 discusses the proposed framework, while Sect. 4 summarizes routing strategies and a comparison methodology. Section 5 provides the results of an application. Finally, a section with discussion and conclusions.

2 Literature Review

Three main modeling and routing approaches are commonly encountered in literature and in practice. In line-based (or service-oriented) networks, container movements are represented as transitions between predefined service loops connecting ports. The temporal dimension is treated in an aggregated manner, with waiting times approximated through average values or frequency-based assumptions [15, 16]. In time-expanded (or diachronic) networks, the temporal dimension is explicitly embedded in the network structure by replicating port nodes for each service call. Arcs represent sailing, waiting, handling, and transshipment operations, ensuring full temporal consistency of routing decisions [17]. Event-based routing algorithms, such as the Connection Scan Algorithm (CSA), operate directly on temporally ordered service connections without requiring the explicit construction of a full time-expanded graph [12, 18]. These algorithms provide exact, time-feasible routing solutions while significantly reducing memory requirements and preprocessing effort.

Each of these approaches involves different trade-offs between modeling accuracy, computational efficiency, and scalability. Line-based models [19, 20] offer compact representations and fast computation, but rely on simplifying assumptions regarding waiting times and schedule synchronization. Time-expanded networks [21, 22] provide a detailed and temporally consistent representation of container movements, but this comes at the expense of a larger network structure. Event-based algorithms, such as CSA, constitute an intermediate paradigm, preserving temporal consistency while avoiding explicit network expansion.

Frequency-based service modeling is widely used in urban public transport to capture the stochastic and adaptive nature of passenger route choices [23–25]. In such systems, passengers can adapt their choices at stops based on service availability, leading to the formulation of hyperpaths that represent sets of feasible routes and associated choice probabilities [26–28]. Hyperpath-based models support stochastic assignment and real-time adaptive behavior. However, in the context of container shipping, the adaptive component at ports is typically not considered. Containers do not make real-time choices analogous to passengers waiting at a stop. Instead, routing decisions are defined deterministically at the strategic or tactical level, producing a single route from the point

of origin to the destination. Consequently, hyperpath formulations are not employed, and container routing is naturally modeled as a shortest-route problem on deterministic networks. This fundamental difference justifies the use of route-based rather than strategy-based representations in liner shipping applications.

Time-expanded network models explicitly incorporate the temporal dimension by replicating nodes over time. Originally developed in the context of public transport scheduling and airline operations, these models have been widely adopted for time-dependent routing and scheduling problems [29–32]. More recently, event-based algorithms, such as CSA, have emerged as efficient alternatives for timetable-based routing, offering exact solutions while avoiding the explicit construction of large time-expanded graphs. An additional motivation for considering temporally consistent routing models lies in the operational reality of liner shipping. Container routing is not a static problem: during service execution, unplanned events such as missed port calls, port congestion, or vessel delays may require re-routing of containers to alternative services or ports [3, 11]. Re-routing involves recomputing a feasible itinerary from the current state of the container onward, potentially modifying the sequence of ports and services while preserving the deterministic nature of the routing decision.

3 Modeling Framework

Given a set of ports P and liner services S, the container routing problem consists of finding minimum-cost routes between all feasible origin-destination port pairs, where costs are represented by travel time and handling time.

3.1 CSA Temporal Network Model

CSA modeling approach adopts a connection-based temporal representation of the liner shipping network and relies on the connection scan algorithm to compute feasible routes. CSA does not explicitly build a graph with temporal nodes and arcs. Instead, it operates directly on an ordered set of connections, exploiting their temporal structure to compute earliest-arrival routes. Let C denote the set of all connections consistent with S in the network. Each connection $c \in C$ is defined as a tuple:

$$c = \left(p_c^{dep}, p_c^{arr}, s_c, t_c^{dep}, t_c^{arr}\right) \tag{1}$$

where:

$p_c^{dep} \in P$ and $p_c^{arr} \in P$ are the departure and arrival port of the connection;
$s_c \in S$ identifies the liner service operating the connection;
t_c^{dep}, t_c^{arr} are the scheduled departure and arrival times.

Given an origin port $o \in P$ and a departure time t_0, CSA computes earliest-arrival times by scanning the ordered list of connections exactly once. For each port p, the algorithm maintains the earliest known arrival time $\tau(p)$, initialized as:

$$\tau(o) = t_0$$

$$\tau(p) = +\infty \forall p \neq o \tag{2}$$

During the scan, a connection c is relaxed if:

$$\tau(p_c^{dep}) \leq t_c^{dep} \tag{3}$$

in which case the arrival time at the destination port is updated as:

$$\tau(p_c^{arr}) \leftarrow \min\left\{\tau(p_c^{arr}), t_c^{arr}\right\} \tag{4}$$

By storing predecessor information, CSA also enables the reconstruction of a time-feasible route between origin and destination ports. In the CSA framework, transshipment is not explicitly modeled, it arises when two consecutive connections share the same arrival and departure port, and the departure time of the second connection is later than the arrival time of the first. Waiting times at ports are implicitly captured by the time difference between the arrival of one connection and the departure of the next feasible one. As a result, both waiting and transshipment feasibility are ensured without introducing additional structural elements. The CSA-based temporal model provides an efficient and schedule-consistent approach to container routing in liner shipping networks. It is particularly well suited for large-scale networks and extensive origin–destination analyses, where the explicit construction of a time-expanded graph may become computationally prohibitive.

3.2 RAPTOR-Based Temporal Routing Model

The modeling approach discussed in this section is based on the Round-Based Public Transit Routing (RAPTOR) algorithm and adopts a schedule-based representation of liner shipping services. Originally developed for public transport routing [12], RAPTOR computes optimal routes by iteratively relaxing services over successive rounds, each corresponding to an additional transfer. This characteristic makes RAPTOR particularly well suited for container routing, where the number of transshipments is a key operational performance indicator. Unlike graph-based formulations, RAPTOR does not explicitly construct a network of nodes and arcs. Instead, it operates directly on the ordered stop sequences of services, thereby avoiding the construction of large time-expanded graphs while preserving full temporal consistency.

Each service $s \in S$ be defined by an ordered sequence of port calls:

$$s = \left(\left(p_1, t_1^{arr}, t_1^{dep}\right), \left(p_2, t_2^{arr}, t_2^{dep}\right), ..., \left(p_{n_s}, t_{n_s}^{arr}, t_{n_s}^{dep}\right)\right), \tag{5}$$

where $p_i \in P$ is the port of the i^{th} call and $t^{\text{arr}}{}_i$, $t^{\text{dep}}{}_i$are the corresponding arrival and departure times.

Each service is assumed to operate according to a fixed schedule within the planned horizon considered. Service frequencies and offsets are handled analogously to the CSA formulation. RAPTOR computes the earliest arrival times by iterating over a sequence of rounds $k = 0, 1, \ldots, K$, where each round corresponds to routes using at most k transshipments. For each port p, the algorithm maintains a label $\tau_k(p)$, representing the earliest arrival time at port p using at most k transfers. Initialization is given as in Eq. 2.

At each round ($k \geq 1$), RAPTOR processes all services that are reachable from ports improved in the previous round ($k - 1$). For a given service s, the algorithm identifies the earliest boardable port call (p_i, $t^{\text{dep}}{}_i$) such that:

$$\tau_{k-1}(p_i) \leq t_i^{dep} \tag{6}$$

From this point onward, the service is scanned forward, and arrival times at subsequent ports are updated as:

$$\tau_k(p_j) \leftarrow \min\left\{ \tau_k(p_j), t_j^{arr} \right\}, \quad j > i \tag{7}$$

This mechanism ensures that temporal feasibility is strictly respected and that each round captures routes with an additional transfer.

The wait times at ports are evaluated through the comparison between arrival times $\tau_{k-1}(p_i)$ and scheduled departure times t_i^{dep}. A container may board a service only if it arrives at the port sufficiently early, thus ensuring temporal causality without the need for explicit waiting arcs. Transshipment corresponds to transitions between rounds. A route that reaches the destination in k rounds is guaranteed to involve at most k transshipment operations. As a result, RAPTOR provides a framework for controlling and analyzing the trade-off between transit time and the number of transfers. By storing predecessor information in each label update, this approach allows the reconstruction of complete routes, including the sequence of ports and services used. This feature enables the derivation of canonical representations consistent with those obtained from the other modeling approach (CSA), facilitating a direct comparison of routing outcomes.

Conceptually, RAPTOR and CSA both operate on schedule-based representations and avoid the explicit construction of time-expanded networks. However, while CSA focuses on earliest-arrival optimization through a single scan of temporally ordered connections, RAPTOR structures the search around the number of transfers. As a consequence, CSA is particularly efficient for computing single optimal routes, whereas RAPTOR supports multi-criteria extensions and the generation of Pareto optimal solutions. This distinction is relevant in the context of container routing, where routing flexibility and the availability of alternative routes are increasingly important under volatile service deployment conditions.

4 Routing and Comparison Methodology

This section describes the routing procedures adopted for each model and formalizes the methodology used to compare the resulting container routes. Given the structural differences between the representations, a common canonical format is introduced to ensure a consistent and meaningful comparison.

4.1 Routing Procedures

For each modeling approach, routing is performed by solving a minimum-cost route problem between all feasible origin-destination port pairs (o,d) $\in$ P $\times$ P, with o $\neq$ d. All

routing computations are based on generalized time-related costs expressed in hours, ensuring comparability between models. Timetable-based routing algorithms (CSA and RAPTOR) provide a relevant methodological perspective for container routing in liner shipping, and the canonical representation provides a common analytical space in which routing outcomes can be compared across different approaches (Table 1).

Table 1. Conceptual comparison of routing paradigms for liner shipping networks

Feature	CSA	RAPTOR
Network representation	Timetable	Timetable
Temporal feasibility	Explicit	Explicit
Primary optimization focus	Earliest arrival	Transfers (rounds)
Handling of waiting times	Explicit	Explicit
Control of transshipments	Indirect	Explicit
Re-routing capability	High	Moderate
Scalability	High	High

Taken together, the comparison highlights that connection-based and round-based timetable algorithms offer complementary perspectives on container routing in liner shipping networks. While CSA is particularly suited for time-feasible routing and disruption management, round-based approaches such as RAPTOR provide a structured framework to analyze transshipment minimization, which becomes increasingly relevant in highly volatile service environments.

4.2 Canonical Representation

To enable a consistent comparison across the different solutions, each computed route must be transformed into a canonical representation (defined as an ordered sequence of port-service pairs). The canonical sequence stores the ordered succession of ports and their associated liner services, independently of the temporal realization of events or the internal structure of the underlying network model. Service changes between consecutive elements of the sequence explicitly identify transshipment events. The canonical representation is derived from the original routes according to model-specific extraction rules, summarized below. In both approaches, the output consists of an ordered sequence of (temporally) feasible service connections. Each connection directly defines a canonical event through its departure port, arrival port, and associated liner service. Note that waiting times and temporal feasibility are implicitly handled by the ordering of connections and are discarded during canonicalization. The resulting canonical sequence is obtained by collapsing consecutive connections belonging to the same service and retaining only effective service legs.

Figure 1 reports an example of the canonicalization process for CSA and RAPTOR. The generated routes are projected onto an equivalent canonical sequence of

port-service pairs, enabling consistent comparison. Canonical representation retains the essential operational decision (the sequence of ports visited, and the liner services used) while abstracting from model-specific constructs such as diversion nodes, waiting arcs, or connection ordering. Therefore, it defines a unified analytical space in which routing outcomes obtained from different modeling paradigms can be compared in a consistent manner. On this basis, the following section introduces the routing procedures and quantitative metrics used to assess similarities and differences between the routes produced by the two approaches.

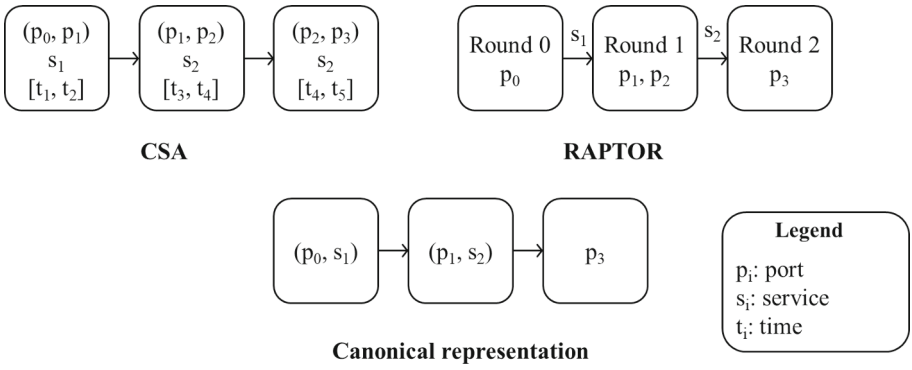


Fig. 1. Routing paradigms and canonical representation

Formally, given a route R between an origin port and a destination port, its canonical representation is defined as follows:

$$R = \left[(p_0, s_1), (p_1, s_2), ..., (p_{k-1}, s_k), p_k\right] \tag{8}$$

where:

- p_0 is the origin port,
- p_k is the destination port,
- $p_i \in P$ are the intermediate ports visited,
- $s_i \in S$ denotes the liner service used to travel from port p_{i-1} to port p_i.

4.3 Similarity Metrics

To quantify the differences between the routing results obtained with different models, several similarity metrics are introduced. All metrics are computed on canonical representations.

Let R^A and R^B be two canonical routes and P^A and P^B the sequences of ports visited. Their similarity is measured using a normalized longest common subsequence (LCS) index:

$$L_p(R^A, R^B) = \frac{LCS(P^A, P^B)}{\max\left(\left|P^A\right|, \left|P^B\right|\right)} \tag{9}$$

The LCS metric is widely used to compare ordered sequences in routing and transportation applications, as it captures structural similarity while allowing for local deviations [33]. Its application to route comparison has been adopted in multimodal routing and network robustness analyses [34].

Similarly, let S^A and S^B denote the corresponding sequences of liner services. Service similarity is defined as:

$$L_s(R^A, R^B) = \frac{LCS(S^A, S^B)}{\max\left(\left|S^A\right|, \left|S^B\right|\right)} \tag{10}$$

This metric focuses on the operational structure of the route, emphasizing the continuity or fragmentation of service usage. Comparing service sequences rather than ports alone is common in public transport and freight network analyses, where service changes represent critical operational events.

The difference in the number of transshipments between two routes is computed as:

$$\Delta T\left(R^A, R^B\right) = T\left(P^A\right) - T\left(P^B\right) \tag{11}$$

The number of transshipments is a key indicator of route quality in liner shipping, as each transshipment entails additional handling, potential delays, and increased operational risk [13, 35]. This metric is therefore commonly used in comparative studies of container routing strategies.

The similarity between the set of route nodes is quantified using the normalized Jaccard index, a standard metric in network analysis to compare set overlap [36, 37]. To capture ordered visitation patterns, the longest common subsequence method is employed, as discussed in related work on route comparison [38, 39]. The node-based Jaccard similarity index measures the overlap between the sets of ports, independently of their order. The Jaccard similarity is defined as:

$$J(R^A, R^B) = \frac{\left|P^A \cap P^B\right|}{\left|P^A \cup P^B\right|} \tag{12}$$

The index takes values in the interval [0, 1], where 1 indicates that the two routes visit exactly the same set of ports. High node-based similarity values indicate agreement on the set of intermediate ports. The LCS metric provides additional information on the order of movements, which is particularly relevant for transshipment routes. Together, these metrics allow for a comparison of different routing solutions.

The comparison analysis is conducted per O/D pair and then aggregated across all considered port pairs. This structured comparison framework allows the identification of systematic differences between the two modeling approaches and provides insight into the conditions under which one representation may be preferred over the other.

5 Application

The objective is to assess the impact of network complexity on routing outcomes and performance under controlled and progressively more complex demanding scenarios. The experiments are carried out in a liner shipping network composed of a set of ports

and regularly scheduled liner services. Each service is defined by an ordered sequence of port calls and associated sailing and port dwell times. A base network configuration is selected as a reference scenario, characterized by a realistic but moderate level of complexity. This base configuration serves as a starting point for generating a family of networks with increasing size and density. All input data are kept identical across the two modeling approaches to ensure comparability of results.

5.1 Scenario Definition

The CMA-CGM [40] shipping lines are considered to define the experimental scenarios. For each service, the ordered sequence of ports calls and indicative sailing times between consecutive calls are available. All services were then converted to a consistent representation using standardized port identifiers and uniform time units; services with incomplete information or ambiguous port sequences were excluded from the analysis. The scenarios of increasing complexity were generated by progressively increasing the number of ports, services, and the length of the planning horizon, as summarized in Table 2. The scenarios are classified into three main categories:

- *Scenario 1* (Small-scale), characterized by a limited number of ports and services, used for qualitative routes comparison and visualization;
- *Scenario 2* (Medium-scale), representative of regional or short-sea shipping networks, used to assess typical routing differences and performance;
- *Scenario 3* (Large-scale), reflecting global liner shipping networks, used to evaluate scalability and potential limitations of the approach.

This classification facilitates the interpretation of results and supports a structured discussion of scalability trade-offs.

Table 2. Classification of experimental scenarios according to network complexity

Scenario ID	Ports	Lines	Services	O/D pairs	Time Horizon [days]
1	19	4	4	342	42
2	47	6	6	2162	105
3	416	310	313	172,640	161

Although this approach allows controlled experimentation, it may not fully capture all sources of uncertainty observed in real-world operations. The use of a public dataset ensures transparency and facilitates the reproducibility of the experimentation.

5.2 Results

Table 3 and Table 4 report a summary of the results. CSA processes a chronologically ordered set of connections, and the number of connections grows with the complexity of the scenario (about 122,000 connections in scenario 3). Preparation times are low (always

below 0.1s except in scenario 3), reflecting the lightweight nature of connection generation. However, scan times increase sharply in scenario 3, where the exhaustive sweep of all connections leads to a computational burden (exceeding 8,600 s). This behavior highlights how CSA maintains memory efficiency but may become time consuming in extreme-scale networks with dense service schedules.

RAPTOR exhibits similar behavior. Calendar generation remains inexpensive, staying under 1.5 s even in scenario 3, confirming the efficiency of its preprocessing stage. Scan times, however, grow with input size: from around 2 s in scenario 1 to over 856,000 s in scenario 3. This underscores RAPTOR's sensitivity to the number of stops and services when applied to dense, large liner networks.

Table 3. CSA: connections and computational effort

	Scenario 1	Scenario 2	Scenario 3
Connections	1,089	4,452	122,088
Prep. Time [s]	0.031	0.046	3.374
Scan Time [s]	0.055	2.740	8,635.360

CSA offers a graph-free, temporally consistent alternative with low memory requirements, though scan times may become prohibitive in very large cases. RAPTOR remains competitive in small and medium instances but encounters substantial performance degradation as network density increases. These contrasts have implications for selecting routing methodologies in increasingly large-scale liner shipping networks.

Table 4. RAPTOR: network size and computational effort for the routing approaches

	Scenario 1	Scenario 2	Scenario 3
Calendar Gen. Time [s]	0.057	0.090	1.464
Scan Time [s]	2.022	310.17	856,387.92

Table 5 summarizes the main routing indicators averaging over all origin-destination pairs for the two procedures in the three scenarios.

Excluding scenario 1, the CSA approach generates longer travel times than RAPTOR. This trend is accompanied by an increase in the number of transshipments: in CSA, the average number and the maximum number of transshipments (from 5.35 to 7.53 and from 2 to 15 respectively) are greater than in RAPTOR (from 2.49 to 4.57 and from 1 to 9 respectively).

The route metrics, based on canonical representation, era reported in Table 6, allowing a model-independent comparison of routing outcomes.

Pairwise comparisons between CSA and RAP confirm that these two temporally oriented algorithms diverge substantially in complex settings. While similarity levels

Table 5. Routing outcomes averaged over all origin-destination pairs

Scenario	Procedure	X_1(*)	X_2(*)	X_3(*)	X_4(*)
Scenario 1	CSA	2760.80	5.35	0.91	2
	RAPTOR	2771.50	2.49	0.49	1
Scenario 2	CSA	2194.90	7.78	2.21	5
	RAPTOR	1789.90	3.15	1.15	3
Scenario 3	CSA	1495.70	7.53	5.56	15
	RAPTOR	1435.80	4.57	2.57	9

(*)X_1, average time in hours; X_2, average number of ports; X_3, average number of transshipments; X_4, maximum number of transshipments.

remain moderate in scenario 1 (Jaccard 0.80), they drop in scenarios 2 and 3, with identical routes falling to 12.30% in scenario 2, and below 1% in scenario 3.

These findings underscore the importance of canonical representations, making such comparisons possible despite the fundamentally different modeling paradigms underlying each routing method.

The similarity analysis lacks in pointing out the overall trade-offs between modeling paradigms when network size, routing outcomes, computational effort, and structural divergence are jointly considered. To synthesize these heterogeneous dimensions, a a weighted average synthesis of the global similarity patterns is reported in Table 6.

Table 6. Similarity metrics based on canonical representation

Scenario	Jaccard	LCS	Identical [%]
Scenario 1	0.80	0.82	16.70
Scenario 2	0.55	0.62	12.30
Scenario 3	0.47	0.53	0.90

6 Discussion and Conclusions

The liner shipping industry is undergoing a profound transformation, marked by increasing operational volatility and frequent reconfiguration of service networks. Recent evidence on record high-capacity churn across major trade lanes confirms that the relative stability characterizing pre-pandemic liner networks has largely disappeared. In this context, the choice of routing model is no longer a purely technical issue, but a substantive modeling decision with direct implications for how network behavior, resilience, and feasibility are represented.

The pairwise comparison between the Connection Scan Algorithm and RAPTOR highlights how their performance profiles respond to varying network conditions, offering valuable insights to support decision-making in liner shipping. Using canonical

representation (that enables model-independent analysis of routing results), this study supports the selection of the most suitable routing algorithm to minimize transit times, reduce the number of transfers, and manage computational resources.

The CSA produces routing outcomes that diverge from RAPTOR, with similarity scores dropping as network complexity increases. CSA routes are longer in terms of travel time, exhibit the highest numbers of transshipments, and differ the most in terms of service and port sequences. Identical-route percentages fall to nearly zero in the largest scenarios for all comparisons. This pattern reflects the algorithm's temporal, earliest-arrival logic: CSA exploits every feasible connection in chronological order, exploring a far richer and more fragmented set of temporal opportunities. As a result, CSA uncovers routes that remain inaccessible to round-based temporal scans (RAPTOR). In fact, RAPTOR produces spatially compact routes, visiting substantially fewer ports. Its multi-round scanning logic prioritizes temporally efficient transfer sequences.

From a modeling standpoint, these findings highlight a spectrum of trade-offs between structural stability and temporal responsiveness:

- CSA emerges as a powerful tool for disruption-aware and operational-level decision making, where fine-grained temporal changes, missed connections, or irregular schedules require rapid re-evaluation of earliest-arrival opportunities;
- RAPTOR offers a balance between temporal realism and spatial compactness, making it suitable for medium-scale operational analyses where fast computation and moderate temporal accuracy are required.

The methodologies and comparative insights presented in this study are of direct interest to several stakeholders. Shipping line operators and network planners can leverage the findings to select routing strategies for minimizing transshipments or improving schedule robustness. Port authorities and terminal managers may take advantage of the approach for managing incoming containers flows. In addition, developers of maritime decision-support systems and logistics technology providers may apply these results to enhance the design of routing engines. Finally, policy makers may find comparative analysis valuable for scenario assessment and long-term strategy development.

Even though this study offers a comprehensive comparison of routing algorithms under controlled scenarios, several limitations should be marked. The analysis is based on deterministic routing, without considering elements such as unexpected delays or schedule reliability. Besides, the procedures may require additional adaptation to specific IT or integration with other systems. Addressing these limitations opens opportunities for future research, including real-time data integration and the exploration of hybrid or probabilistic routing models.

This work raises several perspectives for future research. A first direction concerns hybrid modeling, combining structural stability from line based or time expanded network models with temporal reactivity from CSA or RAPTOR, for example through multi-stage routing or selective temporal expansion. A second direction involves multi-objective formulations, incorporating reliability, buffer times, or service regularity alongside earliest-arrival logic to mitigate the excessive fragmentation observed in CSA-based solutions. Further developments could also explore transfer-penalized connection scanning, priority-aware RAPTOR variants, or probabilistic schedule models to better represent uncertainty.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Russo, F., Musolino, G., Assumma, V.: An integrated procedure to estimate demand flows of maritime container transport at international scale. Int. J. Shipp. Transp. Logist. **6**, 112–132 (2014). https://doi.org/10.1504/IJSTL.2014.059566
2. Russo, F., Pedà, G., Musolino, G.: Attributes influencing port times of container ships. WMU J. Marit. Affairs **23**, 375–392 (2024). https://doi.org/10.1007/s13437-024-00336-7
3. Russo, F., Musolino, G.: Transportation system models to analyse ports competition and cooperation. WMU J. Marit. Affairs **23**, 393–413 (2024). https://doi.org/10.1007/s13437-024-00345-6
4. Musolino, G., Cartisano, A., Fortugno, G.: Structural factors for a third-generation port: planning interventions for mechanical logistics in Gioia Tauro, Italy. WIT Trans. Built Environ. **204**, 55–66 (2021). https://doi.org/10.2495/UT210051
5. Wang, S., Meng, Q.: Schedule design and container routing in liner shipping. Transp. Res. Rec. **2222**, 25–33 (2011). https://doi.org/10.3141/2222-04
6. Wang, S., Meng, Q., Sun, Z.: Container routing in liner shipping. Transp. Res. Part E Logist. Transp. Rev. **49**, 1–7 (2013). https://doi.org/10.1016/j.tre.2012.06.009
7. Baird, A.J.: The economics of container transhipment in northern Europe. Marit. Econ. Logist. **4**, 249–280 (2002). https://doi.org/10.1057/palgrave.ijme.9100048
8. Baird, A.J.: Optimising the container transhipment hub location in northern Europe. J. Transp. Geogr. **14**, 195–214 (2006). https://doi.org/10.1016/j.jtrangeo.2004.12.004
9. Kozan, E., Preston, P.: Genetic algorithms to schedule container transfers at multimodal terminals. Int. Trans. Oper. Res. **6**, 311–329 (1999). https://doi.org/10.1111/j.1475-3995.1999.tb00158.x
10. Alharbi, A., Wang, S., Davy, P.: Schedule design for sustainable container supply chain networks with port time windows. Adv. Eng. Inform. **29**, 322–331 (2015). https://doi.org/10.1016/j.aei.2014.12.001
11. Sea Intelligence: Sea-Intelligence - Press Room. https://www.sea-intelligence.com/press-room/. Accessed 02 Mar 2026
12. Delling, D., Pajor, T., Werneck, R.F.: Round-based public transit routing. Transp. Sci. **49**, 591–604 (2015). https://doi.org/10.1287/trsc.2014.0534
13. Balakrishnan, A., Karsten, C.V.: Container shipping service selection and cargo routing with transshipment limits. Eur. J. Oper. Res. **263**, 652–663 (2017). https://doi.org/10.1016/j.ejor.2017.05.031
14. Castilla-Rodríguez, I., Expósito-Izquierdo, C., Melián-Batista, B., Aguilar, R.M., Moreno-Vega, J.M.: Simulation-optimization for the management of the transshipment operations at maritime container terminals. Expert Syst. Appl. **139**, 112852 (2020). https://doi.org/10.1016/j.eswa.2019.112852
15. Notteboom, T.E.: The time factor in liner shipping services. Marit. Econ. Logist. **8**, 19–39 (2006). https://doi.org/10.1057/palgrave.mel.9100148
16. Wang, T., Wang, S., Meng, Q.: Liner Ship Fleet Planning: Models and Algorithms. Elsevier, Amsterdam (2017)
17. Zheng, J., Hou, X., Qi, J., Yang, L.: Liner ship scheduling with time-dependent port charges. Marit. Policy Manag. **49**, 18–38 (2022). https://doi.org/10.1080/03088839.2020.1849840
18. Dibbelt, J., Pajor, T., Strasser, B., Wagner, D.: Connection scan algorithm. ACM J. Exp. Algorithmics **23**, 1.7:1–1.7:56 (2018). https://doi.org/10.1145/3274661

19. Pan, J.-J., Zhang, Y.-F., Fan, B.: Strengthening container shipping network connectivity during COVID-19: a graph theory approach. Ocean Coast. Manag. **229**, 106338 (2022). https://doi.org/10.1016/j.ocecoaman.2022.106338
20. Jiang, L., Chen, L., Wang, W., Wei, W., Lv, Z., Wang, H.: Advanced network representation learning for container shipping network analysis. IEEE Netw. **35**, 182–187 (2021). https://doi.org/10.1109/MNET.011.2000444
21. Huang, C.: The optimization of empty shipping container reposition based on the network simplex method. In: 2011 International Conference on Management and Service Science, pp. 1–4 (2011). https://doi.org/10.1109/ICMSS.2011.5998249
22. Martin, F., Hemmelmayr, V.C., Wakolbinger, T.: Integrated express shipment service network design with customer choice and endogenous delivery time restrictions. Eur. J. Oper. Res. **294**, 590–603 (2021). https://doi.org/10.1016/j.ejor.2021.02.014
23. Gallo, G., Pallottino, S.: Shortest path methods: a unifying approach. In: Gallo, G., Sandi, C. (eds.) Netflow at Pisa. Mathematical Programming Studies, vol. 26, pp. 38–64. Springer, Heidelberg (1986). https://doi.org/10.1007/BFb0121087
24. Spiess, H., Florian, M.: Optimal strategies: a new assignment model for transit networks. Transp. Res. Part B Methodol. **23**, 83–102 (1989). https://doi.org/10.1016/0191-2615(89)90034-9
25. Nguyen, S., Pallottino, S.: Equilibrium traffic assignment for large scale transit networks. Eur. J. Oper. Res. **37**, 176–186 (1988). https://doi.org/10.1016/0377-2217(88)90327-X
26. Di Gangi, M., Cantarella, G.E., Vitetta, A.: Solving stochastic frequency-based assignment to transit networks with pre-trip/en-route path choice. EURO J. Transp. Logist. **8**, 661–681 (2019). https://doi.org/10.1007/s13676-019-00142-9
27. Nguyen, S., Pallottino, S., Gendreau, M.: Implicit enumeration of hyperpaths in a logit model for transit networks. Transp. Sci. **32**, 54–64 (1998). https://doi.org/10.1287/trsc.32.1.54
28. Nuzzolo, A., Comi, A.: A normative optimal strategy in intelligent transit networks. Transp. Res. Procedia **27**, 380–387 (2017). https://doi.org/10.1016/j.trpro.2017.12.088
29. Cooke, K.L., Halsey, E.: The shortest route through a network with time-dependent internodal transit times. J. Math. Anal. Appl. **14**, 493–498 (1966). https://doi.org/10.1016/0022-247X(66)90009-6
30. Liu, T., (Avi) Ceder, A.: Integrated public transport timetable synchronization and vehicle scheduling with demand assignment: a bi-objective bi-level model using deficit function approach. Transp. Res. Procedia **23**, 341–361 (2017). https://doi.org/10.1016/j.trpro.2017.05.020
31. Liu, T., (Avi) Ceder, A.: Research in public transport vehicle scheduling. In: Handbook of Public Transport Research, pp. 388–408. Edward Elgar Publishing (2021)
32. van Lieshout, R.N., van der Schaft, T.: Dynamic discretization discovery for the multidepot vehicle scheduling problem with trip shifting. INFORMS J. Comput. (2025). https://doi.org/10.1287/ijoc.2024.0698
33. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms. McGraw-Hill (2009)
34. Huang, D., Liu, Z., Jia, Z., Liu, Z.: Customised bus routing problem for commuting travel demand considering route similarity. Transportmetrica A Transp. Sci., 1–28 (2025). https://doi.org/10.1080/23249935.2025.2528960
35. Christiansen, M., Fagerholt, K., Nygreen, B., Ronen, D.: Ship routing and scheduling in the new millennium. Eur. J. Oper. Res. **228**, 467–483 (2013). https://doi.org/10.1016/j.ejor.2012.12.002
36. Jaccard, P.: Nouvelles recherches sur la distribution florale. Bulletin de la Société vaudoise des Sciences Naturelles **44**, 223–270 (1908)

37. Hwang, C.-M., Yang, M.-S., Hung, W.-L.: New similarity measures of intuitionistic fuzzy sets based on the Jaccard index with its application to clustering. Int. J. Intell. Syst. **33**, 1672–1688 (2018). https://doi.org/10.1002/int.21990
38. Chen, L., Özsu, M.T., Oria, V.: Robust and fast similarity search for moving object trajectories. In: Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data, New York, NY, USA, pp. 491–502. Association for Computing Machinery (2005). https://doi.org/10.1145/1066157.1066213
39. Vlachos, M., Kollios, G., Gunopulos, D.: Discovering similar multidimensional trajectories. In: Proceedings 18th International Conference on Data Engineering, pp. 673–684 (2002). https://doi.org/10.1109/ICDE.2002.994784
40. CMA CGM: CMA CGM | Search shipping lines. https://www.cma-cgm.com. Accessed 13 Jan 2025

Analysis and Interpretation of Attributes Affecting Ship Time in Container Ports

Giuseppe Musolino[1(✉)], Paola Panuccio[2], Corrado Rindone[1], Francesco Russo[1], and Antonino Vitetta[1]

[1] DIIES, Università Mediterranea di Reggio Calabria, 89100 Reggio Calabria, Italy
giuseppe.musolino@unirc.it

[2] dAeD, Università Mediterranea di Reggio Calabria, 89100 Reggio Calabria, Italy

Abstract. The ship time in port is considered one of the most representative indicators of the performances of container ports and, more generally, of the countries' competitiveness in the international trade. The work presents the final results of a research project that aims to analyse and interpret the main attributes, or factors, affecting ship times in container ports and to establish correlations between the ship times and specific factors at the national and local (port) levels. The analysis and interpretation of the main attributes was conducted by calibrating descriptive models the put in relation the ship times with different categories of attributes related to the examined ships, ports, maritime services and the geographical areas. The results obtained could support the identification of the determinants that influence the competitiveness of container ports in order to identify and prioritize investments in port infrastructures.

Keywords: commercial ports · ship times in container ports · Transport System Models (TSMs) · attributes interpretation and evaluation · final results

1 Introduction

International trade has historically been based on the maritime transport of goods, which in turn depends on the structure of maritime routes and port systems. Among the most representative indicators of port performance is ship time in port, or ship turnaround time. This variable is one of the key indicators of the performances of container ports and, more generally, of the countries' competitiveness in the international trade [1].

In the above context, a research project titled: "National and local factors affecting times of ships in container ports", financed by the Italian Piano Nazionale di Ripresa e Resilienza (PNRR) inside the NextGenerationEU programme, was carried out by the Department of Information, Infrastructure and Sustainable Energy Engineering (DIIES) of Università Mediterranea di Reggio Calabria (Italy), as principal investigator, and the Department of Engineering of Università di Messina (Italy). The research project aims to identify the main attributes, or factors, affecting ship times in container ports and to establish correlations between the ship times and specific factors at the national and local (port) levels [2]. Ports are generally studied in the scientific literature as individual

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 543–555, 2027.
https://doi.org/10.1007/978-3-032-30530-5_34

nodes, while there is no transversal and integrated knowledge about the set of ports of each country, although quantitative evaluations are necessary to support decisions about national investments in port infrastructures and policies. The existing research focuses mainly on single ports or terminals [3, 4], with a limited number of studies among different ports [5–7].

The methodologies adopted in the research project find their theoretical foundation in the framework of Transport System Models (TSMs) (see [8–14]). TSMs are conceived to simulate the behavior of both travelers and freight transport operators across multiple choice dimensions when interacting with transport infrastructures and services. TSMs consist of three main components: supply models, travel demand models, and assignment models. Supply models estimate the transport costs associated with the use of infrastructures and services. Supply models support the study the centrality that each port, or terminal, has in the structure of international maritime services [5, 15, 16]. Travel demand models estimate the choices of users based on the transport costs. The path choice models, in particular, estimate users' choices among available path alternatives [17–19]. Supply-demand interaction, or assignment, models calculate the flow and the (congested) costs on the elements of the network from the interaction between the multiple user's choices and the transport costs [11, 12, 20]. The literature about freight transport modelling concerns two levels: commodity-based and vehicle-based models, as reported in [21, 22] and [23].

The paper presents a three-level statistical analysis of the determinants of ship times in container ports, grounded on country-level macroeconomic, infrastructural and logistical indicators as well as port-level data. Statistical-descriptive models are employed to establish correlations between the average ship time in container ports, considered a dependent variable, with a set of attributes, as independent variables, defined at the national and port levels. The development of models is mainly based on the acquisition of data derived from Automatic Identification System (AIS) [24, 25]. As an example, Belcore and Polimeni [26, 27] introduced a modular framework to assess the phases of a ship call at port, allowing the identification of anchoring and berth operations, as well as the estimation of dwell time and waiting time. Such an approach also allowed to assess the performances of a port by means of a set of spatial-temporal indicators (spatial density, average service time, berth use) to investigate the operations performed in container ports [28].

Ship time in container ports may be considered an important indicator for assessing the competitiveness of countries and ports in the global supply chain, providing insights for the definition of intervention priorities and of appropriate measures to improve the performances of freight flow management in container ports and terminals. The results are important as they could support port authorities and planners to build national-level benchmarks among ports and to identify of each port's relative positioning.

The paper, after the introduction, is articulated as follows. Section 2 reports the specification of the model adopted. Section 3 presents the analysis and interpretation of the main attributes that are correlated with ship time in container ports at country and port levels. The last section is dedicated to the conclusions and the future research.

2 Models

The models used in the research belong to the class of statistical-descriptive, as they put in relation the average ship times in container ports, as dependent variable, with different attributes, as independent variables, at country and local levels [29–33]. The relationship is supposed to be one-way.

The general functional form of the model is:

$$t = t(\mathbf{x}; \boldsymbol{\beta}) + e \tag{1}$$

with:

t, ship time in port;
x, vector of attributes, or factors;
β, vector of parameters to be calibrated;
$t()$, function;
ε, error term.

By assuming that $t()$ is linear, Eq. (1) is specified as follows:

$$t = \beta_0 + \Sigma_j \beta_j x_j + \varepsilon \tag{2}$$

where:

x_j is the generic attribute;
β_0 and β_j are components of vector **β**.

The vector of parameters, **β**, may be calibrated by means of the Least Squares (LS) method, which allows to identify the vector $\boldsymbol{\beta}_{LS}$, among the different configurations of vector **β**, that minimizes the sum of the squares of the deviations between the observed values of ship times and the values of ship time calculated with Eq. (2).

The method has been applied both at country and port levels (see Fig. 1). At country level, the observed ship times, the estimated ones with Eq. (2) and the attributes are average values related to each country. At port level, the observed ship times and the estimated ones with Eq. (2) are AIS-derived related to each individual container ship, while the attributes are related to each ship and to the port.

3 Factors Affecting Ship Times in Container Ports

The research aimed to verify the correlation between the ship times of container ports and different attributes, of factors, at country and port levels.

At country-level, the analysis is focused on the performance of the container port system of a country as a whole, while at port-level the analysis concerned each container port.

The sources of data are different:

- country-level: the values of attributes are obtained from different data-base referred to the year 2019: UNCTADSTAT (www.unctadstat.unctad.org); The World Bank (www.worldbank.org); The Global Economy (www.theglobaleconomy.com);

- port-level: the attributes are built ad-hoc on the basis of the database provided by available AIS data [34] related to each individual ship extracted from the open-access platforms vesselfinder (vesselfinder.com) and ships.jobmarineman (ships.jobmarineman.com).

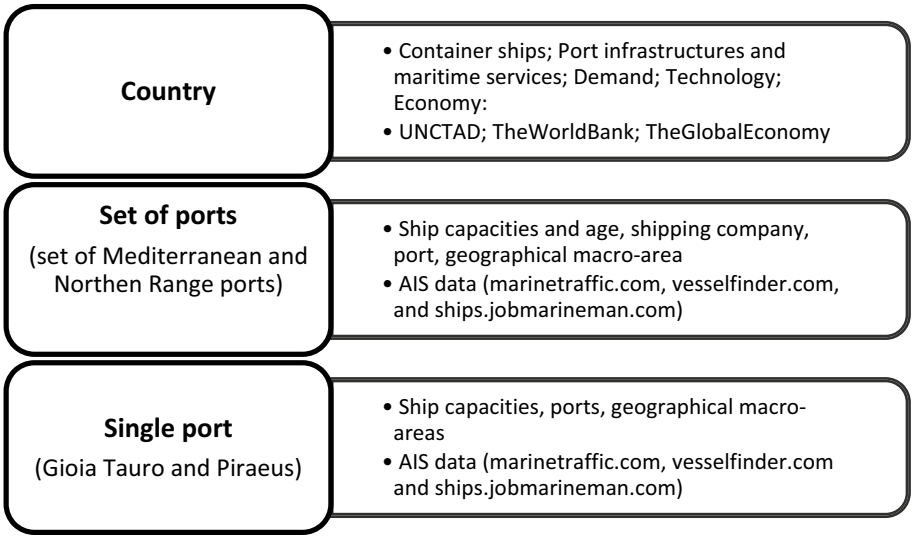


Fig. 1. Categories of attributes and sources of data at each level of analysis

3.1 Country Level

The experimentation at country level was conducted on a set of fifty countries belonging to Europe, Africa, Middle East and far East (see Fig. 2).

Definition of Attributes

The attributes considered in the analysis at country-level can be aggregated into different classes reported below.

The attributes related to the container ships are:

- the average gross tonnage of container ships, which indicates the average gross tonnage of container ships using country's ports during the period considered (year);
- the maximum gross tonnage of container ships, which indicates the maximum gross tonnage of the container ship that used at least one time one of the ports of the country during the period considered (year);
- the average capacity of container ships, which indicates the average capacity of by container ships using country's ports during the period considered (year).
- the maximum capacity of container ships, which is the maximum number of containers carried by a container ship using country's ports during the period considered (year).

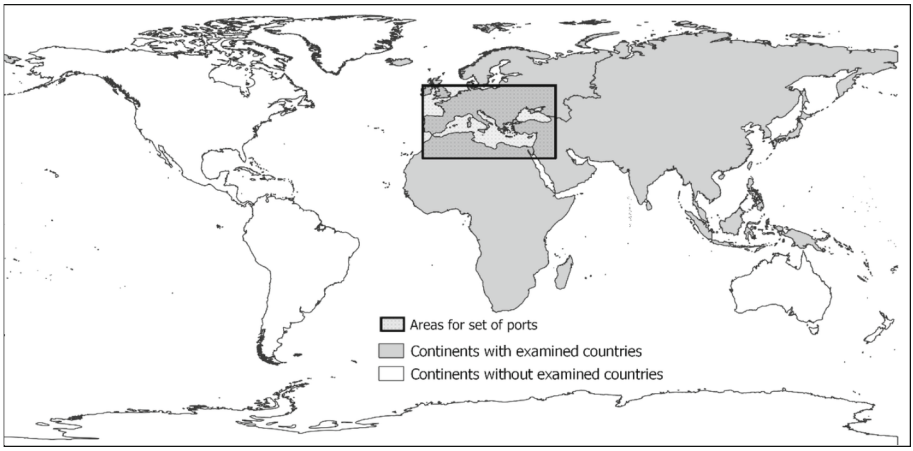


Fig. 2. Continents with examined countries.

The attributes related to the port infrastructures and maritime services are:

- the linear shipping connectivity index, which indicates the level of integration of a country inside the global maritime transport network;
- the logistics performance index, which indicates the efficiency of a country of handling goods both inside and outside the country;
- the port infrastructure index, which assesses the quality of a country's port infrastructure over a reference period (year);
- the number of ship calls, which indicates the number of port calls during the examined period by all ships (e.g. container ships) in all ports of the country.

The attribute related to the demand is the container throughput, which indicates the number of handled containers in a country during the period considered (year).

Two are the attributes connected to technology:

- the technology-overall index, which estimates technological capabilities of a country in relation to physical investment, human capital and technological effort and represents national capacities to use, adopt and adapt new-generation technologies;
- the ICT technology index, which estimates technological capabilities in a similar way to the technology-overall index, but only with respect to ICT technologies.

The attributes related to economy are:

- the gross domestic product, which indicates the average annual growth as a percentage of the country;
- the gross domestic product per capita, which indicates the average annual growth of gross domestic product per capita as a percentage.

Calibration Results

The correlation analysis shows that the ship time in port is highly correlated with the maximum capacity of container ships, the ICT technological index and the number of ship calls. The following general elements emerge:

- the ship time in port presents a declining trend and a reducing dispersion of values for increasing values of maximum capacity of container ships and of technological index;
- the ship time in port rapidly declines for increasing number of ship calls, until the threshold value of 5000 ship/year followed by a slight declining trend.

Several models have been calibrated considering different combinations of attributes belonging to the categories of attributes defined. Some comments concerning some attributes are reported below.

The models specified with the attribute maximum capacity of container ships generally present a negative value of parameter. This indicates that ship time in port reduces when the country is equipped with container ports able to handle ships with higher capacity; that is, countries having ports with larger physical infrastructures (e.g. docks, seabeds, …).

The models specified with the attribute number of ship calls also presents a negative value of parameter. The interpretation could be that an increase of the number of ship calls is associated with a better organization of the country's port system, leading to lower average ship times in port.

As for the two previous attributes, the models specified with the attribute ICT technological index present a negative value of the parameter. In other words, ship time in port reduces as the country's ICT dotation increases.

The ratio between the values of parameters associated with the attributes ICT technological index and maximum capacity of container ships attributes is 7.5. This means that an increase of +10% of the ICT technological index corresponds to an average increase of 750 [TEU] of the maximum capacity of container ships that can potentially be hosted in the ports of a country. Similarly, the ratio between the values of parameters associated with the attributes ICT technological index and number of ship calls is 450 on average. Therefore, an increase +10% of the ICT technological index corresponds to an increase of 450 [ships/year] ship calls in the ports of a country.

The models specified with the attributes linear shipping connectivity index, and logistics performance index, have both positive signs. They may indicate that container ports of countries with more connections and more logistics services are more congested; therefore, container ships spend more time in ports.

3.2 Set of Ports

The attributes of the model refer to a set of ports belonging to two macro-regions:

- Mediterranean Sea: Gioia Tauro, Piraeus, Livorno, Genoa, Valencia, Algeciras and Tanger Med;
- Northern Range: Rotterdam, Hamburg and Antwerp.

The following average values of ship times are calculated from the AIS-derived data in the set of examined ports. The average value of ship time in the port of Gioia Tauro is 34.0 h, while the average values in the ports of Algeciras and Tangier Med are respectively 21.2 h and 23.1 h. The port of Piraeus and Genova presents a value of ship time similar to the one of Gioia Tauro: 32.7 h in Piraeus and 35.0 h in Genova. The

average value is 28.3 h in the port of Rotterdam, while it is 53.1 h in the port of Hamburg and 46.4 h in the port of Antwerp.

Definition of Attributes

The attributes considered in the analysis at port-level can be aggregated into different classes reported below.

Ship:

- Age, [years], age of container ship calculated as difference between the current year and the year of building;
- Five specifications of ship capacity [TEUs] are considered:
  - Capacity1 $\in [0, +\infty]$
  - Capacity2 $\in [0, 12{,}000[$
  - Capacity3 $\in [12{,}000; +\infty]$
  - Capacity4 $\in [12{,}000; 18.000[$
  - Capacity5 $\in [18{,}000; +\infty]$

Shipping company (considering the alliances in year 2022):

- 2M $\in [0, 1]$
- Ocean Alliance $\in [0, 1]$
- The Alliance $\in [0, 1]$,
- Other $\in [0, 1]$

Port:

- Throughput
- PCS $\in [0, 1]$

Geographical macro-area to which the selected ports belong:

- Africa $\in [0, 1]$
- Southern Europe $\in [0, 1]$
- Northern Europe $\in [0, 1]$

Calibration Results

Among the models calibrated considering different combinations of attributes belonging to above categories, a selection of model results is reported below.

Some models have a positive value of calibrated parameter associated to the attribute Capacity1, which indicates that ship time in container port increases as the port is able to handle container ships with higher capacity. This result should be connected to the longer handling times due the higher number of unloading/loading operations of higher capacity container ships. The same consideration is valid for parameter of attribute, Age, which means that ship time in port decreases if the ship that arrives at the container port belongs to more recent generations. This could be due to the fact that ships built more recently have on-board equipment that allow more efficient operations.

A group of models has been specified with the attributes Capacity2 and Capacity3 in order to highlight possible non-linearity effects in the capacity attribute (2-class capacity). The calibrated parameters show that the ship time in ports increases as the ship

capacity increases, as for models specified with the attribute Capacity1. In particular, the ratio between the parameters associated to Capacity 2 and Capacity3 is 2/3; this means that container ships that have a capacity higher than 12,000 TEU has a greater impact on the ship times in port.

Another group of models is specified with the attributes associated to the ship capacity Capacity2, Capacity4 and Capacity5 with the same objective of the previous group of models (3-class capacity). It is worth noting that the value of parameter associated to the attribute Capacity4 is greater than the values of parameters associated the other two classes: Capacity2and Capacity5. This means that higher ship times in port are associated to the class of container ships with capacity between 12,000 and 18,000 TEUs. The lower value of parameter associated to attribute Capacity5 could be justified by the fact that the latest generation of ships (ULCS and New Panamax types) have priority over the ships with less capacity in the container ports considered.

Models specified with geographical attributes, as Southern Europe and Northern Europe, present positive calibrated parameters. This indicates that container ports of Southern Europe and Northern Europe have higher ship times than African ports (Tangier Med). The result is important because it establishes some reference thresholds of ship times in container ports belonging to the three examined geographical macro-areas.

3.3 Single Port: Gioia Tauro and Piraeus

The paragraph describes a further analysis at port-level concerning two ports: Gioia Tauro (Italy) and Piraeus (Greece). The container terminal of Gioia Tauro is managed by Terminal Investment Limited (TIL belonging to MSC Group). It is located along the Suez-Gibraltar route with deep sea-beds up to 18 m and 1.7 km of quays extension. TIL has recently renewed the fleet of vehicles inside the yard by switching from straddle carriers to transtainers, allowing containers to be stacked from two to three-four levels. As consequence, the port increased the handling and storing capacity and the container throughput passed to approximately 3.9 million of TEUs in 2024 and 4.5 million of TEUs in 2025. Container traffic by rail mode with Northern Italy and Central Europe is less developed than the large hub ports of Northern Europe. The gateway handled almost 1,000 trains/year in 2025. The large hinterland area has strong development potentials; but the initiatives of the regional Calabrian government of creating a location-based and sector-based special economic zone (Calabria Region 2016, 2018 [35, 36]) were modified by the Italian national government by creating a special economic zone for the entire South of Italy and covering the whole economic sectors (Musolino et al. 2023 [37]; Russo and Musolino, 2023 [38]). The container terminals of port of Piraeus are operated by the Piraeus Container Terminal company, under the control of COSCO Group. The container throughput oscillated between 4.0 and 5.0 million of TEUs in the last years. COSCO Group established in Piraeus its hub for routes and commercial strategies in the Mediterranean Sea. The port has no rail and road connections with Central Europe (Piraeus Port Authority, 2024 [39]; Galani, 2025 [40]).

Definition of Attributes

The following attributes are considered at this stage.

Ship:

- Age, [years];
- Five specifications of ship capacity [TEUs] are considered:

  – Capacity1 $\in [0, +\infty]$
  – Capacity2 $\in [0, 5{,}000[$
  – Capacity3 $\in [5{,}000; +\infty]$
  – Capacity4 $\in [5{,}000; 15{,}000[$
  – Capacity5 $\in [15{,}000; +\infty]$

Port:

- Reference Port $\in [0; 1]$, equal to 1 if the port is Gioia Tauro or Piraeus, 0 otherwise:
- Distance from Previous Port call $\in [0, +\infty]$, distance between the previous port of call and the reference port (Gioia Tauro or Piraeus);
- Distance from Next Port call $\in [0, +\infty]$, distance between the reference port (Gioia Tauro or Piraeus);
- Transhipment Previous Port $\in [0, 1]$, is equal to 1 if when the previous port called is a transhipment hub port, otherwise 0:
- Transhipment Next Port $\in [0, 1]$, is equal to 1 if when the next port called is a transhipment hub port, otherwise 0:

Geographical Area

- Previous Port call $\in [0, 1]$, it is equal to 1 if the previous port of call is located in one of the three geographical areas: Indo-Pacific, Mediterranean, North Sea; it is equal to 0 if the previous port of call is located in the geographical area of Atlantic;
- Next Port call $\in [0, 1]$, it is equal to 1 if the next port of call is located in one of the three geographical areas: Indo-Pacific, Mediterranean, North Sea; it is equal to 0 if the next port of call is located in the geographical area of Atlantic;
- Chain $\in [0, 1]$ is equal to 1 if the previous and next port call belong to the Mediterranean area, otherwise it is 0.

Parameters' Calibration

Several specifications of models have been calibrated. Some comments of the main results are reported below.

Preliminary, the parameter defines the reference threshold of ship port time for Gioia Tauro and Piraeus. The value of the calibrated parameter of the attribute Reference Port is 2.37 [hours], as expected, confirming the differences in the observed average values of ship times of the two ports.

Models specified with the attributes Capacity2 and Capacity4 show that ships with capacity lower than 5,000 TEU present an increasing ship time than the ships with capacity between 5,000 and 15,000 TEU. This result is in line with what happens in the transhipment hub container ports, like Gioia Tauro and Piraeus, where mother ships have priority in the port respect to feeder ships.

Models specified with attributes belonging to the category Geographical area show that ships arriving from ports belonging to the macro-area of the Indo-Pacific, the Mediterranean and the North Sea (or having a port as next destination belonging to

the macro-area of the Indo-Pacific, the Mediterranean or the North Sea), have a lower port time than Atlantic, assumed as reference area.

4 Conclusions

Container ports are complex systems, where three main categories of elements coexist and concur to their competitiveness:

a) physical infrastructures on the seaside, like breakwaters, quays, draught; and on the landside, like loading/unloading areas, last-mile connections with terrestrial networks; warehouses;
b) port governance and port organization, such as maritime services operated by shipping companies, handling and storage services operated by terminal companies;
c) immaterial infrastructures, such as emerging ICT, research and education.

The paper presents an analysis and interpretation of the main attributes, or factors, affecting ship time in container ports, which is considered one of the most representative indicators of the performances of container ports and, more generally, of the countries' competitiveness in the international trade.

The analysis has been conducted at country and port levels. The analysis and interpretation of the attributes is executed by calibrating descriptive models that put in relation the ship time in container ports with different categories of attributes related to the ships, the ports, the maritime services and the geographical areas examined.

The proposed three-level structure of analysis allows to obtain results ranging from the macro level, country, to the micro level, port. The finding that ICT investments and physical infrastructure investments may have equivalent marginal effects on ship time could be a result of practical significance for port planners. The selection of Gioia Tauro and Piraeus as a case study is motivated as these ports are considered to be major transhipment hub ports in the Mediterranean area.

The limitations are reported in the following. The temporal scope and the undiscussed sensitivity of the results regarding the identification of thresholds of the classes of container ship capacity constrain the generalisability and reproducibility of the findings. The data at the country level concern the year 2019, a pre pandemic period, which do not allow to discuss whether the identified relationships are expected to hold under the altered shipping conditions belonging to the post 2020 period, characterized by severe port congestion, supply chain disruptions, accelerated digital transformation, war in Ukraine.

Further developments concern the investigation about the influence of further attributes not considered in the models presented and the extension of the data-base by adding observations about new container ports and post-2020 years.

Acknowledgements. This study is carried out within the research project "National and local factors affecting times of ships in container ports", Piano Nazionale di Ripresa e Resilienza (PNRR), Next Generation EU, Progetti di Rilevante Interesse Nazionale PRIN_2022_PNRR_P202292YFW, CUP: C53D2300878001. This presentation reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

References

1. World Bank Port Reform Toolkit. World Bank, Washington, DC (2025)
2. Musolino, G., Panuccio, P., Rindone, C., Russo, F., Vitetta, A.: National and local factors affecting ship time in container ports: system of models at country and port levels. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15897, pp. 144–155. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97660-5_11. ISBN 978-3-031-97659-9
3. Tsou, M.-C.: Discovering knowledge from AIS database for application in VTS. J. Navig. **63**, 449–469 (2010). https://doi.org/10.1017/S0373463310000135
4. Mazibuko, D.F., Mutombo, K., Kuroshi, L.: An evaluation of the relationship between ship turnaround time and key port performance indicators: a case study of a southern African port. WMU J. Marit. Affairs **23**, 499–524 (2024). https://doi.org/10.1007/s13437-024-00330-z
5. Ducruet, C.: The geography of maritime networks: a critical review. J. Transp. Geogr. **88**, 102824 (2020). https://doi.org/10.1016/j.jtrangeo.2020.102824
6. Ducruet, C., Itoh, H.: Spatial network analysis of container port operations: the case of ship turnaround times. Netw. Spat. Econ. (2022). https://doi.org/10.1007/s11067-022-09570-z
7. Campisi, T., Russo, A., Trwdy, E., Zanne, M., Tesoriere, G.: The importance of the centrality of ports for passenger transport in the Adriatic-Ionian Basin. In: Gervasi, O., Murgante, B., Garau, C., Taniar, D., Rocha, A.M.A.C., Faginas Lago, M.N. (eds.) ICCSA 2024. LNCS, vol. 14823, pp. 269–282. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-65329-2_18. ISBN 978-3-031-65328-5
8. Ben-Akiva, M.E., Lerman, S.R.: Discrete Choice Analysis: Theory and Application to Travel Demand; MIT Press Series in Transportation Studies. MIT Press, Cambridge (1985). ISBN 978-0-262-02217-0
9. Cascetta, E., Nuzzolo, A., Biggiero, L., Russo, F.: Passenger and freight demand models for the Italian transportation system. In: Proceedings of the 7th World Conference on Transport Research, vol. II
10. de Dios Ortúzar, J., Willumsen, L.G.: Modelling Transport, 3rd edn. Wiley, Chichester, New York (2001). ISBN 978-0-471-86110-2
11. Cascetta, E.: Transportation Systems Engineering Theory and Methods. Springer, New York (2013). ISBN 978-1-4757-6873-2
12. Russo, F., Musolino, G.: Transportation system models to analyse ports competition and cooperation. WMU J. Marit. Affairs **23**, 393–413 (2024). https://doi.org/10.1007/s13437-024-00345-6
13. Musolino, G., Vitetta, A.: Modelling the energy consumption of road vehicles in transport networks evaluation. Int. J. Energy Res. **2025**, 5226552 (2025). https://doi.org/10.1155/er/5226552
14. Russo, F., Musolino, G., Assumma, V.: An integrated procedure to estimate demand flows of maritime container transport at international scale. IJSTL **6**, 112 (2014). https://doi.org/10.1504/IJSTL.2014.059566
15. Russo, A., Campisi, T., Bouhouras, E., Basbas, S., Tesoriere, G.: Sustainable maritime passenger transport: a network analysis approach on a national basis. In: Gervasi, O., et al. (eds.) ICCSA 2023. LNCS, vol. 14109, pp. 195–207. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-37120-2_13. ISBN 978-3-031-37119-6
16. Rindone, C., Russo, A.: Centrality analysis for railway network planning: constrained measures for high speed. Transp. Eng. **22**, 100393 (2025). https://doi.org/10.1016/j.treng.2025.100393
17. Quattrone, A., Vitetta, A.: Random and fuzzy utility models for road route choice. Transp. Res. Part E Logist. Transp. Rev. **47**, 1126–1139 (2011). https://doi.org/10.1016/j.tre.2011.04.007

18. Vitetta, A.: A quantum utility model for route choice in transport systems. Travel Behav. Soc. **3**, 29–37 (2016). https://doi.org/10.1016/j.tbs.2015.07.003
19. Vitetta, A.: Advancing path choice in transport systems: insights from fuzzy logic models. Sustainability **18**, 3236 (2026). https://doi.org/10.3390/su18073236
20. Cantarella, G.E.: Dynamics and Stochasticity in Transportation Systems: Tools for Transportation Network Modelling. Elsevier, Amsterdam (2019). ISBN 978-0-12-814353-7
21. Ogden, K.W.: Urban Goods Movement: A Guide to Policy and Planning. Ashgate, Aldershot (1992). ISBN 978-1-85742-029-6
22. De Jong, G., Vierth, I., Tavasszy, L., Ben-Akiva, M.: Recent developments in national and international freight transport models within Europe. Transportation **40**, 347–371 (2013). https://doi.org/10.1007/s11116-012-9422-9
23. Tavasszy, L., de Jong, G.: Modelling Freight Transport. Elsevier (2014). ISBN 978-0-12-410400-6
24. Polimeni, A., Belcore, O.M.: Times of ship in container ports: AIS data for maritime transport and ports applications (2024)
25. Rindone, C.: AIS data for building a transport maritime network: a pilot study in the strait of Messina (Italy). In: Gervasi, O., Murgante, B., Garau, C., Taniar, D., Rocha, A.M.A.C., Faginas Lago, M.N. (eds.) ICCSA 2024. LNCS, vol. 14823, pp. 213–226. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-65329-2_14. ISBN 978-3- 031-65328-5
26. Belcore, O.M., Di Gangi, M., Polimeni, A.: Automatic identification system data to assess container port performances. Transp. Res. Procedia **86**, 612–619 (2025). https://doi.org/10.1016/j.trpro.2025.04.077
27. Belcore, O.M., Di Gangi, M., Polimeni, A.: Open AIS data to assess seaside operations in container ports. In: Crupi, V., Bucci, V. (eds.) Progress in Marine Science and Technology. IOS Press (2025). ISBN 978-1-64368-610-3
28. Belcore, O.M., Polimeni, A.: Marine transport analyses: a toolkit to assess vessel operations at the container port. In: Proceedings of the 2025 IEEE International Conference on Environment and Electrical Engineering and 2025 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe), Chania, Crete, Greece, July 15 2025, pp. 1–7. IEEE (2025)
29. Russo, F., Peda, G., Musolino, G.: Container ports in country systems: calibration of the aggregate function for the time of the ship in port. Int. J. TDI **6**, 415–427 (2022). https://doi.org/10.2495/TDI-V6-N4-415-427
30. Russo, F., Pedà, G., Musolino, G.: Attributes influencing port times of container ships. WMU J. Marit. Affairs **23**, 375–392 (2024). https://doi.org/10.1007/s13437-024-00336-7
31. Musolino, G., Pedà, G., Russo, F.: ICT factors affecting ship times in container ports: experimental analysis at country level. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15897, pp. 131–143. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97660-5_10. ISBN 978-3-031-97659-9
32. Russo, F., Musolino, G.: The role of emerging ICT in the ports: increasing utilities according to shared decisions. Front. Future Transp. **2**, 722812 (2021). https://doi.org/10.3389/ffutr.2021.722812
33. Russo, F., Rindone, C.: Container maritime transport on an international scale: data envelopment analysis for transhipment port. WIT Trans. Ecol. Environ. **150**, 831–843 (2011)
34. Belcore, O.M., Polimeni, A.: Times of ships in container ports: automatic identification system data for analyzing traffic conditions at a maritime terminal. Front. Future Transp. **6**, 1735788 (2026). https://doi.org/10.3389/ffutr.2025.1735788
35. Calabria Region Piano Regionale Dei Trasporti 2016
36. Calabria Region Strategic Development Plan of Special Economic Zone (2016)

37. Musolino, G., Cartisano, A., Fortugno, G.: Structural factors for a third-generation port: planning interventions for mechanical logistics in Gioia Tauro, Italy. In: Proceedings of the Urban and Maritime Transport 2021, WIT Transactions on the Built Environment, vol. 204, 16 June 2021, pp. 55–66 (2021)
38. Musolino, G., Cartisano, A., Fortugno, G.: Special economic zones planning for sustainable ports: aggregate economic impact of port of Gioia Tauro. In: Gervasi, O., Murgante, B., Misra, S., Rocha, A.M.A.C., Garau, C. (eds.) ICCSA 2022. LNCS, vol. 13381, pp. 60–71. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-10548-7_5. ISBN 978-3-031- 10547-0
39. Piraeus Port Authority Full-Year Results (2024)
40. Morgan, K.: The Routledge History of the Modern Maritime World Since 1500, 1st edn. Routledge, London (2025). ISBN 978-1-00-360691-8

Modal Choice for High-Speed Railway Systems: The Italian Case Study

Giuseppe Musolino, Marialuisa Moschella, Francesco Russo(✉), and Domenico Sgro

DIIES, Università Mediterranea of Reggio Calabria, 89100 Reggio Calabria, Italy
francesco.russo@unirc.it

Abstract. High-Speed Railway (HSR) systems have established worldwide for their technical characteristics, which wherever they are built have allowed greater economic, social and environmental sustainability of transport. 60,000 km of HSR in operation and 60,000 km under construction or design testify to the increasing importance of HSR systems. For the planning of new lines and for the monitoring of the current ones, it is necessary to have knowledge of modal choice that allow us to evaluate the goodness of the infrastructure investments and the validity of the services offered to users. To this end, it is essential to collect data on users' choices in the current situation.

The overall results obtained provide valuable insights for understanding current transport flows and provide a basis for analysing future trends resulting from the introduction of new HSR lines in southern Italy.

The effectiveness of the HSR system depends on the consistency between technical planning and design choices and international standards. Verifying this consistency is made possible by an in-depth analysis of demand, carried out using mobility data. The preliminary results obtained represent a first step in this direction.

Keywords: High-Speed Rail (HSR) · Sustainability · Demand · Modal Choice

1 Introduction

In the context of the 2030 Agenda's sustainable development goals, the planning of high-speed rail (HSR) infrastructure is a key factor to achieving more sustainable mobility, with implications on the socio-economic aspects, increasing the economic competitiveness of the connected regions, and on environmental aspects by promoting a more energy-efficient mode of transport contributing to the European Green Deal's target of reducing transport sector emissions by 2050 [1–11]. In specific regions such as Greece, southern Italy and Eastern Europe, where per capita GDP is below the European average, high-speed rail lines have been recognised as infrastructure solutions for reducing inequalities caused by limited connectivity with Europe's major economic hubs. For this reason, high-speed rail (HSR) plays a central role in the planning of the Trans-European Transport Network (TEN-T), with Regulation (EU) 2024/1679, which defines the strategic corridors and priority projects that ensure balanced connectivity across Europe [12, 13].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 556–569, 2027.
https://doi.org/10.1007/978-3-032-30530-5_35

1.1 Operational Models for HSR

The main factor that makes HSR infrastructure strategically important is the drastic reduction in travel times between the origins and destinations served. Indeed, travel time is considered the key factor determining competitiveness between high-speed rail and air for distances of up to around 1,000 km [14]. This competitiveness is closely linked to the type of operating model adopted. According to the classification by Campos and De Rus (2009), there are four operating models for rail services which define how the HSR service interacts with the existing infrastructure [15] (Fig. 1):

- Model 1 is characterised by separate infrastructures for high-speed and conventional trains and is known as the "Exclusive exploitation model";
- Model 2 refers to railway infrastructures where high-speed trains run on both dedicated lines and conventional sections, and is known as the "Mixed high-speed model";
- Model 3 provides for certain conventional trains to be able to use high-speed lines, this is known as the "Mixed conventional model";
- Model 4 allows conventional and high-speed trains to run on any infrastructure and is known as the "Fully mixed conventional model".

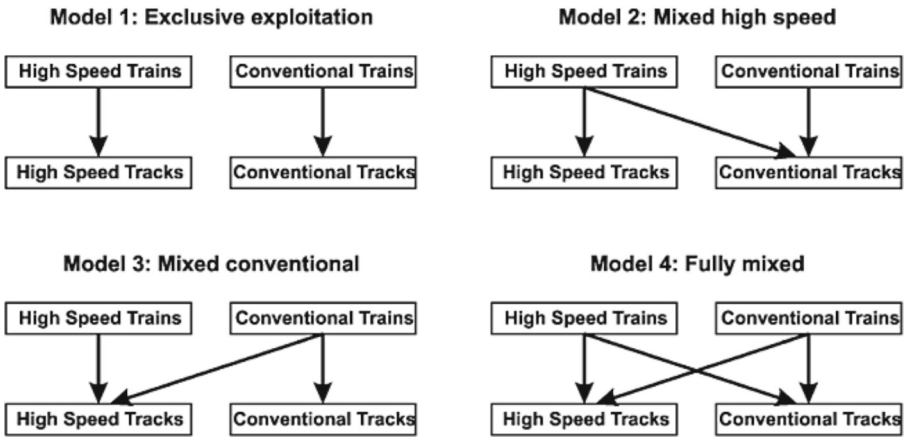


Fig. 1. HSR operational models [15].

Model 1 is the only dedicated service, whilst the other three belong to the mixed-service category. This distinction is significant because the competitiveness of the HSR service is reflected in the transition from mixed services to dedicated services, and thus to Model 1. Therefore, the dedicated lines generate the reduction in HSR train travel times which ensure the impacts on centrality and territorial attractiveness [16].

The improvement in the level of service associated with Model 1 (exclusive) has a direct effect on transport demand. According to the relevant scientific literature, there are specific time thresholds that influence user choices: for journeys with travel times of less than 2,5 h, the market is dominated by the HSR train; for travel times of more than

2,5 h and less than 3,5 h it is equally divided between the HSR train and air mode; for travel times of 3,5 to 5 h HSR still play a significant contribution; finally, above 5 h the air mode is dominant [14].

1.2 HSR Demand Analysis

As mentioned above, the analysis of travel demand is fundamental to understanding changes in passenger behaviour in relation to the transition from the various mixed operating Models (2, 3 and 4) to the exclusive operating Model 1. The travel demand generated by the introduction of a dedicated HSR line comprises three components that manifest themselves at different time intervals [17].

The first segment of demand consists of demand that shifts from other modes or transport services to the HSR network, known as diverted demand. This modal shift is determined by travel time and specific characteristics of the alternative modes, such as fares and the frequency of services offered. This component becomes apparent during operating time.

The second segment of demand relates to the new passenger flows, this is known as induced demand. The increase in demand is driven by the rise in the frequency of journeys made by the same users who were already travelling before the introduction of the dedicated HSR line and is linked to the significant reduction in travel times. The impact of this component on demand is a result of tactical choice.

The third segment of demand is demand driven by economic growth. This component is linked to new travel flows resulting from the economic and social impacts that the presence of the HSR line has on the region over time. In this case, the increase in demand is not directly linked to the improvement in service but is the result of structural changes in the territory. The impact of this component on demand occurs as a result of strategic outcomes.

The three components of demand can be estimated using Transport System Models (TSM) [18–21], in particular, mode/service choice models are used to estimate direct demand [22], while travel frequency models and destination choice models are used to estimate induced demand [23, 24], models capable of quantifying the relationships between economic development and mobility, known as SETI (Spatial Economic Transport Interaction), to estimate demand arising from economic growth [25, 26].

The application of these models depends on the availability of data. The scientific literature uses two types of surveys to analyse diverted and induced demand: Revealed Preference (RP) surveys, which reveal users' actual behaviour within the currently operating transport system, and Stated Preference (SP) surveys, which present realistic but non-operational infrastructure scenarios. In some situations, the RP approach can be replaced by the acquisition of datasets provided by agencies that monitor mobility flows.

1.3 LARG Paradigm

Analysing demand for HSR services therefore requires a methodological approach capable of describing users' choice behaviour and identifying the different segments of overall demand. To pursue this endpoint, the LARG paradigm provides a conceptual framework

for the development of new HSR lines that takes into account both strategic planning and technical design [2, 5, 8, 10, 11]. The LARG paradigm is based on four dimensions: *lean* dimension aims to reduce unnecessary costs without compromising service quality; *agile* dimension focuses on the infrastructure's capacity to respond to transport demand, estimating the three components of demand described above as factors determining service effectiveness; *resilient* dimension, which assesses the railway system's ability to respond to risk events such as safety and service regularity by restoring normal operating conditions; *green* dimension, which analyses the environmental sustainability of the infrastructure project by comparing energy consumption and pollutant emissions with other modes of transport, and in particular air mode.

This work falls within the LARG paradigm, focusing specifically on the *agile* dimension, which is dedicated to the study of demand, applying it to the Italian context, where the strategic role of extending HSR network has been identified as one of the priority challenges in the Piano Nazionale di Ripresa e Resilienza (PNRR) to reduce economic and infrastructural disparities between the northern and southern Italy [27].

This study aims to analyse the current modal split for travel between southern Italy and Rome, in both directions, with a particular focus on the Reggio Calabria–Rome relationship, which is currently served by a mixed rail service (operational Model 2). Hence, a methodological approach using mobility data derived from telephone datasets is adopted.

This study is useful to policy-makers and planners as it provides a methodological basis for analysing the impact of an HSR line on transport demand.

Following this introduction, the paper comprises Sect. 2, which explains the methodology for estimating demand; Sect. 3, which outlines the experimental study; and Sect. 4, which consists of the discussion and the final comments.

2 Methodology for Demand Estimation

2.1 Data Acquisitions

Estimating transport demand for high-speed services requires a dataset that reflects both users' actual choice behaviour within the current infrastructure context and their future behaviour, simulated by exploring infrastructure scenarios that have not yet been implemented. Thus, questionnaires are designed to investigate users' Revealed Preferences (RP) for the current situation and questionnaires designed to investigate users' Stated Preferences (SP) with regard to one or more hypothetical scenarios.

The questions contained in an RP questionnaire relate to a specific origin-destination (O-D) pair are designed to gather information on departure and arrival times, the reason for the journey, the mode of transport chosen for the journey, the frequency of the same journey, and any basic socio-economic information such as occupation, level of education, and the number of household members. As an integration to the RP questionnaire, data on specific origin-destination relationships can be obtained using nationwide mobility databases. These databases are provided by companies collecting mobility data and

storing information on observed passenger flows; they may include mode of travel, journey times, the average number of trips on O-D pairs for representative samples of travellers, and socio-economic information. The data contained in these databases, however, may be subject to anonymisation in accordance with current privacy regulations.

SP questionnaires describe one or more hypothetical scenarios relating to the construction of new infrastructure, which differ in terms of certain key attributes, such as fare levels and travel times for specific O-D pairs. The information gathered through this type of survey makes it possible to estimate changes in user behaviour following the introduction of transport options not yet available in the area. The SP questionnaire investigates two components of passenger demand: diverted demand and induced demand. In particular, the presentation of more than one hypothetical scenario to each respondent allows to collect a large number of observations.

Finally, RP data reflect actual behaviour, obtained from direct questionnaires or mobility databases; SP data, on the other hand, reflect users' simulated behaviour in relation to hypothetical scenarios, conditioned by the reliability of the responses, which may be influenced by the gap between the hypothetical scenarios and the real behaviour users will adopt in the future.

2.2 Models for Demand Estimation

Current and future demand is estimated using TSMs, which also include the models used to analyse supply and the interaction between supply and demand [19, 20, 22, 23, 28–35]. In academic literature, demand analysis refers to the four-stage model system, which breaks down the demand function into the product of subsets relating to different dimensions of choice [20]. The system consists of:

- emission model, which estimates the number of journeys generated by each origin;
- distribution model, which allocates journeys across different destinations;
- modal choice model, which allocates journeys across different modes of transport;
- path choice model, which defines the use of links in a transport network.

The design of one of these models involves three stages: specification of the functional form; calibration, which assigns values to the parameters based on the data collected from surveys; and validation through informal and statistical tests that verify the model's ability to reproduce reality.

The mathematical models found in the literature can be segmented into two categories, which differ according to the approach used.

The first category comprises models that employ a descriptive statistical approach. These models generate an aggregate estimation of travel demand, which is useful for large-scale analyses or preliminary studies.

The second category of models is based on discrete choice theory [36, 37] and includes models that adopt a probabilistic-behavioural approach to simulate users' choice behaviour. In these models the demand estimation is disaggregated, showing users' behaviour with respect to different variables.

Belonging to the second category are RUM (Random Utility Model) models, QUM (Quantum Utility Model) models and FUM (Fuzzy Utility Model) models [36–51]. RUM models are extensively explored in literature and are based on the assumption that the

user chooses the alternative that maximises utility. Utility consists of a deterministic component and a stochastic component that captures measurement errors. QUM and FUM models are recent extensions trying to overcome the limitations of RUM models.

This work has been carried out as part of the MOST Project – Spoke 4, which focuses on analysing both the demand for and supply of high-speed rail services [6, 22, 23, 30, 32, 33, 35, 52–55].

2.3 The Database Structure

The RP data has been acquired through the Vodafone Business Analytics service, which provides non-personal databases generated by the mobile network. The dataset is built in full compliance with privacy protection regulations, the information has been anonymised directly at source and disseminated only in aggregated form. The characteristics of the dataset used for this study are explained in Table 1.

Table 1. Summary of the dataset's features.

Features	Description
Modes of transport investigated	Air, train, bus
Minimum distance threshold for displacement detection	Only travel distances of more than 100 km
Number of travels	Aggregate count of travels by combination of attributes
Classification of passengers	Resident of origin, resident of destination, non-resident, foreign user

The "number of travels" information expresses the intensity of mobility flows based on spatial, temporal and modal characteristics. The "number of travels" is not referred to any individual user, but it is an aggregate count of all the journeys recorded during the observation period (35 days, from 15 March to 18 April 2025). Each trip corresponds to a complete sequence of movement that meets the methodological criteria defined by the project:

- True-Origin;
- Origin node (departure terminal);
- Destination node (arrival terminal);
- True-Destination.

The data is obtained by aggregating journeys that share the same combination of spatial (Table 2) and temporal (Table 3) attributes, for example: true-Origin, origin node, destination node, true-Destination, mode of travel, time slot of the journey, day of the week, passenger class. Therefore, this is not a direct count of passengers, but rather an aggregate number of journey that respects a minimum threshold of 15 users per journey. This measure allows to analyze passenger flows.

Table 2. Spatial attributes.

Attribute	Description	Specification
Origin node/ Destination node	Departure or arrival terminal (airport, railway station, hub)	10 airports 21 railway stations 19 bus hubs
True-Origin True-Destination	User's last stop (≥4 h) before departing from the origin infrastructure node User's first stop (≥4 h) after arriving at the destination infrastructure node	Calabria and Sicily: • Region • Province • Municipality Rome, Milan and Naples: • Region • Municipal district • Municipality
Geographical levels for defining travel Home-based	The user is a resident of the area in which the departure or arrival node is located	Region Province/Municipal district Municipality

Table 3. Temporal attributes.

Attribute	Description	Specification
Observation period	35 days from 15 March 2025 to 18 April 2025	Standard period: 15 March 2025 to 4 April 2025 Easter period: 5 April 2025 to 18 April 2025
Day of the week	Representative day on which mobility flow aggregates are measured	Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday
Time slots (24 intervals)	One-hour interval during which the departure from the origin node is recorded	From 00:00–00:59 to 23:00–23:59
Macro-time slots (5 intervals)	Aggregation of time slots into 5 broader intervals	00:00–04:59 05:00–09:59 10:00–15:59 16:00–19:59 20:00–23:59

The data are derived from the population of mobile phone users affiliated with Vodafone company, which has been expanded using inference models to represent the entire traveller universe. The extrapolation process is based on calibration factors that include the distribution of market shares by geographical region and age group, the number of devices associated with each user, and mobile phone usage patterns. To ensure the reliability of the estimates, the model is constantly updated and validated through periodic

comparisons with data observed in known situations, such as sporting events, trade fairs or airports, thus enabling a high level of accuracy to be maintained.

3 Experimental Case Study

This paper focuses on analysing users' choice behaviour in relation to the introduction of a HSR line between Reggio Calabria and Salerno in southern Italy. Currently, the Italian rail network comprises 921 km of operational lines, located mainly in the north of the country [13, 56]. The transport options available for connecting the capital, Rome, with southern Italy include buses, cars, planes, conventional trains and a hybrid service that combines high-speed trains on dedicated lines from Rome to Salerno (operational model 1) and high-speed trains running on conventional lines from Salerno to Reggio Calabria (operational model 2) [57]. The experimental hypothesis behind the research is the implementation of a dedicated line for HSR services on the Salerno–Reggio Calabria section, with the aim of establishing a continuous high-speed service. The analysis of current demand is carried out using nationwide travel data obtained from a database. Future demand is analysed using data collected from a survey conducted at stations and airports in Sicily, Calabria and Rome.

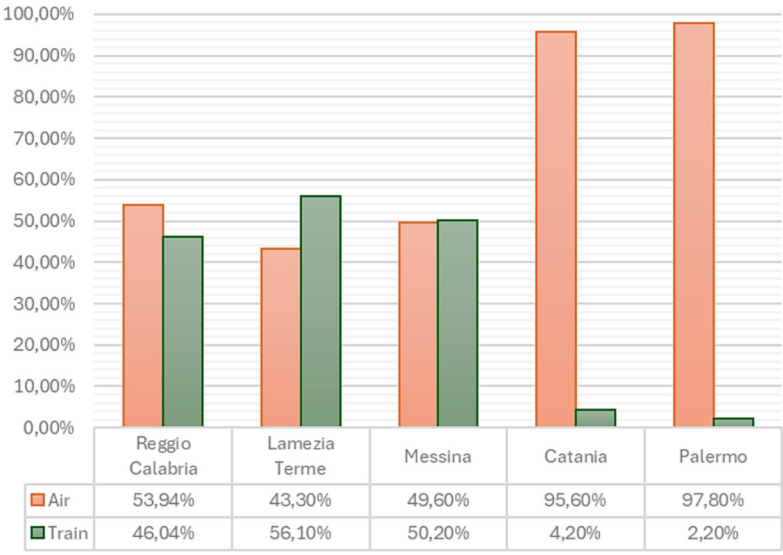
This paper takes an in-depth look at the methodology for collecting and analysing data relating to the current situation (RP).

To ensure privacy protection and statistical significance, the 'number of travels' is provided for aggregations containing at least 15 users. If a combination of attributes comprises fewer than 15 distinct users, the data is not released in its original form but is subject to a progressive masking procedure.

The following graphs show the modal split derived from RP data for travel to Rome (Fig. 2) from the Calabrian cities (Reggio Calabria and Lamezia Terme) and Sicilian cities (Messina, Catania and Palermo) covered by the study, and vice versa (Fig. 3), obtained in accordance with the graph shown as an illustrative example of travel choice between Calabria and Rome (Fig. 4).

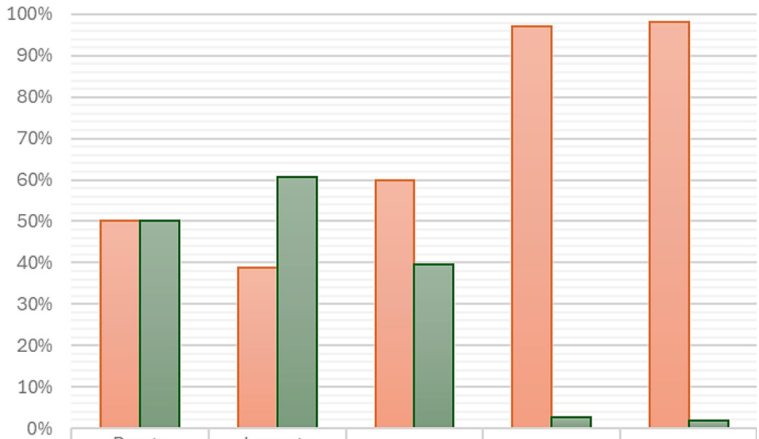
The data collected on the current situation (RP) provide a detailed overview of the modal split for travel between the main cities of Calabria and Sicily and Rome, in both directions. The analysis reveals a clear polarisation in transport choices, with air travel dominating for travel from Sicily.

The results obtained provide the basis for analysing current demand and for applying probabilistic-behavioural demand models specific to TSMs, with the aim of simulating user behaviour in relation to future scenarios concerning the introduction of a dedicated high-speed line.



	Reggio Calabria	Lamezia Terme	Messina	Catania	Palermo
Air	53,94%	43,30%	49,60%	95,60%	97,80%
Train	46,04%	56,10%	50,20%	4,20%	2,20%

Fig. 2. Modal split for travel from Calabria and Sicily to Rome.



	Reggio Calabria	Lamezia Terme	Messina	Catania	Palermo
Air	50%	38,80%	60%	97,20%	98,20%
Train	50%	60,80%	39,60%	2,60%	1,78%

Fig. 3. Modal split for travel from Rome to Calabria and Sicily.

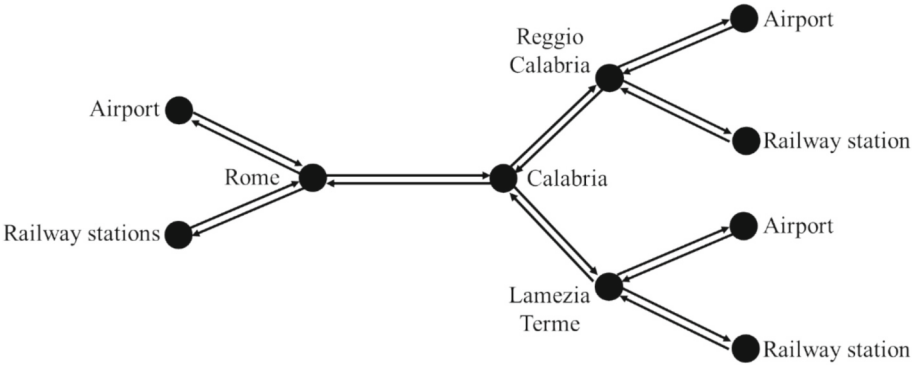


Fig. 4. Graph for the Calabria – Rome connection.

4 Conclusions

This research work falls within the LARG paradigm, focusing on the *agile* dimension dedicated to the study of demand. In order to apply demand models, data must be available. The literature identifies two types of surveys for analysing diverted and induced demand: Revealed Preference (RP) surveys and Stated Preference (SP) surveys. In some cases, the RP approach can be replaced by the acquisition of datasets provided by agencies that monitor mobility flows at national level.

This paper describes a dataset for the collection of RP data on a national scale in Italy. The dataset has been provided by Vodafone Business Analytics and has been constructed in accordance with the research project's guidelines. The database contains anonymised and aggregated data covering a 35-day observation period (15 March–18 April 2025), and includes information on 10 airport nodes, 21 railway nodes and 19 bus hubs, with a minimum distance threshold of 100 km for travel. True-Origin and True-Destination locations have been identified based on a stationary threshold of 4 consecutive hours, allowing the entire mobility chain to be reconstructed.

In contrast to studies based on traditional RP surveys, which are typically limited to small samples and information collected only at infrastructure nodes, this study uses a nationwide mobility dataset covering 35 consecutive days. Furthermore, the integration of "True-Origin" and "True-Destination" allows for the reconstruction of the entire mobility chain (access time, main journey, egress time), overcoming the common limitation of RP data limited to infrastructure nodes.

The preliminary analysis of data identified the current modal split for travels made from the cities of Reggio Calabria and Lamezia Terme, located in Calabria, and from Messina, Catania and Palermo, located in Sicily, to Rome and vice versa.

The results observed can be interpreted in light of discrete choice theory. According to RUM models, the probability of choosing a mode decreases as the associated disutility increases. Each of the attributes identified through the RP data has a specific theoretical expectation regarding its causal effect on mode choice. In particular, travel time, access time, and egress time are attributes of disutility, corresponding to a negative coefficient in the systematic utility function.

The RP data obtained from the dataset reveal patterns consistent with theoretical expectations. For example, the higher share of train travel from Lamezia Terme compared to Reggio Calabria is due to shorter travel time, which reduces the overall disutility of rail travel. Similar, the strong preference for air travel from Catania and Palermo is a result of the high disutility associated with rail travel and the time attribute. This evidence confirms the central role of travel time as a key factor, in line with the literature, on mode choice for medium- to long-distance travel.

These results are useful for studying current demand and for applying stochastic models belonging to the TSM class.

Future research plans to integrate SP data which, together with RP data, will enable the analysis of future transport demand generated by the introduction of a HSR line in a territory, particularly in southern Italy.

Acknowledgements. This study was carried out within the MOST – Sustainable Mobility National Research Center and received funding from the European Union Next-Generation EU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1033 17/06/2022, CN00000023). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

References

1. Santilli, G.: Italia divisa in due dalla Tav: Pil a +7% nelle città collegate (2020). https://www.ilsole24ore.com/art/italia-divisa-due-tav-pil-7percento-citta-collegate-ACCZKIFB
2. Russo, F.: Which high-speed rail? LARG approach between plan and design. Future Transp. **1**, 202–226 (2021). https://doi.org/10.3390/futuretransp1020013
3. European Commission: Sustainable and Smart Mobility Strategy – putting European transport on track for the future (2021). https://transport.ec.europa.eu/document/download/be22d311-4a07-4c29-8b72-d6d255846069_en?filename=2021-mobility-strategy-and-action-plan.pdf
4. Raicu, S., Costescu, D., Popa, M., Dragu, V.: Dynamic intercorrelations between transport/traffic infrastructures and territorial systems: from economic growth to sustainable development. Sustainability **13**, 11951 (2021). https://doi.org/10.3390/su132111951
5. He, S., Mei, L., Wang, L.: The dynamic influence of high-speed rail on the spatial structure of economic networks and the underlying mechanisms in Northeastern China. IJGI **10**, 776 (2021). https://doi.org/10.3390/ijgi10110776
6. Praticò, F.G., Fedele, R.: Economic sustainability of high-speed and high-capacity railways. Sustainability **15**, 725 (2022). https://doi.org/10.3390/su15010725
7. Ji, C., Yao, Y., Duan, J., Li, W.: Sustainable mechanism of the entrusted transportation management mode on high-speed rail and the impact of COVID-19: a case study of the Beijing-shanghai high-speed rail. Sustainability **14**, 1171 (2022). https://doi.org/10.3390/su14031171
8. Li, D., Jiang, R., Lu, Z., Sun, S., Wang, L.: Does the construction of high-speed rail change the development of regional finance? Sustainability **15**, 10641 (2023). https://doi.org/10.3390/su151310641
9. European Environment Agency: National emissions reported to the UNFCCC and to the EU Greenhouse Gas Monitoring Mechanism, April 2024 (2024). https://sdi.eea.europa.eu/catalogue/srv/api/records/6331f651-8863-4656-a911-669f2a332a1e?language=all. https://doi.org/10.2909/6331F651-8863-4656-A911-669F2A332A1E

10. Prakongwittaya, S., Liangrokapart, J.: Integrating LARG measures to improve supply chain transparency performance. Asia Pac. Manag. Rev. **30**, 100324 (2025). https://doi.org/10.1016/j.apmrv.2024.09.002
11. Surmařová, S., Neumannová, M., Pařil, V., Vrána, M., Chmelík, J.: Do I really like to shift to rail? Influence of rail modernisation on passenger preferences. Res. Transp. Bus. Manag. **59**, 101284 (2025). https://doi.org/10.1016/j.rtbm.2024.101284
12. European Union: Regulation (EU) No 1315/2013 of the European Parliament and of the Council (2013)
13. European Union: Regulation (EU) 2024/1679 of the European Parliament and of the Council (2024)
14. Givoni, M., Dobruszkes, F.: A review of ex-post evidence for mode substitution and induced demand following the introduction of high-speed rail. Transp. Rev. **33**, 720–742 (2013). https://doi.org/10.1080/01441647.2013.853707
15. Campos, J., De Rus, G.: Some stylized facts about high-speed rail: a review of HSR experiences around the world. Transp. Policy **16**, 19–28 (2009). https://doi.org/10.1016/j.tranpol.2009.02.008
16. Rindone, C., Russo, A.: Centrality analysis for railway network planning: constrained measures for high speed. Transp. Eng. **22**, 100393 (2025). https://doi.org/10.1016/j.treng.2025.100393
17. Ben-Akiva, M., Cascetta, E., Coppola, P., Papola, P., Velardi, V.: High speed rail demand forecasting in a competitive market: the Italian case study (2010)
18. Nuzzolo, A., Russo, F.: Departure time and path choice models for intercity transit assignment. Pergamon (1998)
19. Ortúzar, J.D.D., Willumsen, L.G.: Modelling Transport. Wiley (2011). https://doi.org/10.1002/9781119993308
20. Cascetta, E.: Transportation Systems Engineering: Theory and Methods. Springer, New York (2013)
21. Russo, F., Rindone, C.: Regional transport plans: from direction role denied to common rules identified. Sustainability **13**, 9052 (2021). https://doi.org/10.3390/su13169052
22. Russo, F., Sgro, D., Musolino, G.: Sustainable development of railway corridors: methods and models for high speed rail (HSR) demand analysis. In: Gervasi, O., et al. (eds.) ICCSA 2023. LNCS, vol. 14110, pp. 527–538. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-37123-3_36
23. Russo, F., Moschella, M., Musolino, G.: Railway demand evaluation: hsr induced component. In: Gervasi, O., Murgante, B., Garau, C., Taniar, D., Rocha, A.M.A.C., Faginas Lago, M.N. (eds.) ICCSA 2024. LNCS, vol. 14822, pp. 173–187. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-65318-6_12
24. Russo, F., Musolino, G., Moschella, M.: High-speed rail travel demand analysis: pilot experimentation on generation model. Transp. Res. Procedia **90**, 822–829 (2025). https://doi.org/10.1016/j.trpro.2025.06.074
25. Cascetta, E., Marzano, V., Papola, A.: Multi-regional input-output models for freight demand simulation at a national level. In: Ben-Akiva, M., Meersman, H., and Van De Voorde, E. (eds.) Recent Developments in Transport Modelling, pp. 93–116. Emerald Group Publishing Limited (2008). https://doi.org/10.1108/9781786359537-006
26. Russo, F., Musolino, G.: A unifying modelling framework to simulate the spatial economic transport interaction process at urban and national scales. J. Transp. Geogr. **24**, 189–197 (2012). https://doi.org/10.1016/j.jtrangeo.2012.02.003
27. CPI, Council Presidency of Italy: Piano Nazionale di Ripresa e Resilienza (PNRR) (2021). https://www.governo.it/sites/governo.it/files/PNRR_0.pdf

28. Russo, F., Rindone, C.: Container maritime transport on an international scale: data envelopment analysis for transhipment port. Presented at the Sustainable Development And Planning 2011 July 12 (2011). https://doi.org/10.2495/SDP110691
29. Russo, F., Musolino, G.: The role of emerging ICT in the ports: increasing utilities according to shared decisions. Front. Future Transp. **2**, 722812 (2021). https://doi.org/10.3389/ffutr.2021.722812
30. Di Gangi, M., Russo, F.: Potentiality of rail networks: integrated services on conventional and high-speed lines. Presented at the COMPRAIL, 2022 November 30 (2022). https://doi.org/10.2495/CR220091
31. Tesoriere, G., Russo, A., De Cet, G., Vianello, C., Campisi, T.: The centrality of Italian airports before and after the COVID-19 period: what happened? Eur. Transp. Trasporti Europei, 1–16 (2023). https://doi.org/10.48295/ET.2023.93.2
32. Russo, F., Sgro, D., Musolino, G.: Dynamic structure of fares for high speed rail services. In: Gervasi, O., Murgante, B., Garau, C., Taniar, D., C. Rocha, A.M.A., Faginas Lago, M.N. (eds.) ICCSA 2024. LNCS, vol. 14822, pp. 188–201. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-65318-6_13
33. Rindone, C., Russo, A.: A network analysis for HSR services in the South of Italy. In: Gervasi, O., Murgante, B., Garau, C., Taniar, D., C. Rocha, A.M.A., Faginas Lago, M.N. (eds.) ICCSA 2024. LNCS, vol. 14822, pp. 217–232. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-65318-6_15
34. Campisi, T., Russo, A., Trwdy, E., Zanne, M., Tesoriere, G.: Some considerations on the analysis of port centrality in the Adriatic basin for tourist transport purposes. Transp. Res. Procedia **83**, 125–132 (2025). https://doi.org/10.1016/j.trpro.2025.02.018
35. Moschella, M.: Domanda passeggeri High Speed Rail (HSR) in aree in ritardo di sviluppo: analisi sperimentale per la relazione Roma-Reggio Calabria. Presented at the Approfondimenti sulla mobilità terrestre (2025)
36. McFadden, D.: The measurement of urban travel demand. J. Public Econ. **3**, 303–328 (1974). https://doi.org/10.1016/0047-2727(74)90003-6
37. Ben-Akiva, M.E., Lerman, S.R.: Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press, Cambridge (1985)
38. Anderson, S.P., De Palma, A., Thisse, J.-F.: Discrete choice theory of product differentiation (1992). https://doi.org/10.7551/mitpress/2450.001.0001
39. Tversky, A., Shafir, E.: The disjunction effect in choice under uncertainty. Psychol. Sci. **3**, 305–309 (1992)
40. Russo, F.: Fuzzy theory in transportation field: fuzzy sets for simulating path choice behavior, pp. 279–283 (1997)
41. McFadden, D.: Disaggregate behavioral travel demand's RUM side - a 30-year retrospective (2001)
42. Cascetta, E., Papola, A.: Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand. Transp. Res. Part C Emerg. Technol. **9**, 249–263 (2001). https://doi.org/10.1016/S0968-090X(00)00036-X
43. Train, K.E.: Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge (2009)
44. Quattrone, A., Vitetta, A.: Random and fuzzy utility models for road route choice. Transp. Res. Part E Logist. Transp. Rev. **47**, 1126–1139 (2011). https://doi.org/10.1016/j.tre.2011.04.007
45. Busemeyer, J.R., Bruza, P.D.: Quantum Models of Cognition and Decision. Cambridge University Press (2012). https://doi.org/10.1017/CBO9780511997716
46. Vitetta, A.: A quantum utility model for route choice in transport systems. Travel Behav. Soc. **3**, 29–37 (2016). https://doi.org/10.1016/j.tbs.2015.07.003

47. Tinessa, F., Papola, A., Marzano, V.: The importance of choosing appropriate random utility models in complex choice contexts. In: 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), Naples, Italy, pp. 884–888. IEEE (2017). https://doi.org/10.1109/MTITS.2017.8005638
48. Papola, A., Tinessa, F., Marzano, V.: Application of the combination of random utility models (CoRUM) to route choice. Transp. Res. Part B Methodol. **111**, 304–326 (2018). https://doi.org/10.1016/j.trb.2018.03.014
49. Ben-Akiva, M., McFadden, D., Train, K.: Foundations of stated preference elicitation: consumer behavior and choice-based conjoint analysis. FNT Econ. **10**, 1–144 (2019). https://doi.org/10.1561/0800000036
50. Croissant, Y.: Estimation of random utility models in R: the mlogit Package. J. Stat. Soft. **95** (2020). https://doi.org/10.18637/jss.v095.i11
51. Vitetta, A.: Advancing path choice in transport systems: insights from fuzzy logic models. Sustainability **18**, 3236 (2026). https://doi.org/10.3390/su18073236
52. Di Gangi, M., Russo, F.: Design of hybrid rail services on conventional and high-speed lines. IJTDI **7**, 113–121 (2023). https://doi.org/10.18280/ijtdi.070206
53. Giunta, M.: Trends and challenges in railway sustainability: the state of the art regarding measures, strategies, and assessment tools. Sustainability **15**, 16632 (2023). https://doi.org/10.3390/su152416632
54. Giunta, M., Leonardi, G.: Framework for life cycle railway sustainability assessment: a methodological approach based on advanced methods and tools. In: Gervasi, O., Murgante, B., Garau, C., Taniar, D., C. Rocha, A.M.A., Faginas Lago, M.N. (eds.) ICCSA 2024. LNCS, vol. 14822, pp. 233–244. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-65318-6_16
55. Panuccio, P.: State of the art of sustainable development of railway nodes: the high speed rail stations in Italy. In: International Conference on Computational Science and Its Applications-ICCSA 2024 (2024, submitted to)
56. UIC: High-Speed Rail Atlas 4 Edition. International Union of Railways (UIC) - Paris 2023 (2023)
57. Russo, F., Moschella, M.: Mode choice at national level: revealed and stated preference survey for high-speed rail realization. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15894, pp. 359–374. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97648-3_24

Environmental Sustainability of Railway Corridors: Case Studies of Italian High Speed Rail

Giuseppe Musolino(✉), Marialuisa Moschella, Francesco Russo, and Domenico Sgro

DIIES, Università degli Studi Mediterranea di Reggio Calabria, Reggio Calabria, Italy
giuseppe.musolino@unirc.it

Abstract. The increase of carbon footprint and the climate change scenario are central themes of the current technical and political debate. An important contribution is provided by the transport of passengers and goods. Over long national distances the problem becomes even more crucial. In a nutshell, the carbon footprint of High-Speed Rail (HSR) is on average, lower than other modes of transport, such as private cars or airplanes, for the operating phases. The reduction of carbon emissions with the use of HSR for passengers' trips at national scale would make it possible to pursue the Sustainable Development Goals (SDGs) included in the 2030 Agenda, with specific reference to Goal 13: "Climate change". It is necessary to carry out an analysis of the current scientific literature on this topic, considering the levels of environmental impacts in relation to the operative phases of the HSR, through an analysis and comparison of the values of the indicators. The purpose of the paper is to give a synthetic answer to the problem: how sustainable are the HSR systems in terms of carbon footprint?

The results will be twofold. On the one hand, parametric values can be defined which can synthetically represent the environmental advantage provided by the implementation of the HSR for passengers' trips; on the other hand, the fields of unknowledge in which it is useful to carry out research can be delimited. The reference result is given by applying the synthetic parametric approach to a case study: the relationship Rome-Milan (Italy). The link is presented in relation to the current airline connections and passengers flow, and considering several scenarios connected with the implementation of the HSR line. The conclusions allow decision-makers and planners to be supported in the design and construction of new HSR lines with a broader point of view, which considers environmental CO_2 emissions.

Keywords: CO2 emissions · environmental impact · Italian case studies · Transport System Model · High Speed Rail

1 Introduction

High-Speed Rail (HSR) mode-service results to be the most innovative terrestrial passenger transport system designed and built in the last decades. HSR has experienced rapid growth both in supply, in terms of services and infrastructures, and demand sides, in

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 570–584, 2027.
https://doi.org/10.1007/978-3-032-30530-5_36

terms of passengers transported. There are 64,698 km of HSR lines in operation, 19,927 km under construction and 53,201 km planned [1]. Specifically, China results to have the largest HSR network with about 45,390 km, whereas the whole Europe continent results to have 12,454 km of lines, the Middle East 1,681 km, the North America 735 km and the Africa content 186 km [1].

It is stated that Italy is one of the first countries which have planned, designed and built a HSR infrastructure. The preparation of the NextGenerationEU financial program specified as Piano Nazionale di Ripresa e Resilienza (PNRR) in Italy [2], has given a significant boost to the development of transport planning on a national scale. Taking into account the impact of carbon footprint, in terms of GreenHouse Gas (GHG), it is well known that HSR in the operating phases is, on average, lower than other transport modes, such as private cars or airplanes. In the year 2021, according to [3], the global emissions of the transport sector results to be about 30%, with China as the main contributor to this statistic, and then the USA and Europe. The impact due to the transport sector depends on the structure of supply and demand, in particular with regard to supply, the main papers are identified which are [4–19] on the other hand, the reference papers of supply are [20–40].

A significant contribution is represented by the evaluation of travel demand [41, 42], with regard to the modes of transport utilized for travels on national scale (e.g., rail, bus, car, air), hence in the users' choices [43]. For these travels it is important to evaluate the environmental impact, in order to provide mitigation solutions [44]. The evaluation is crucial in order to know the alternatives with less environmental impact.

The latter theme plays a crucial role in the actual political and technical discussions at global level, in particular considering its global mitigation and reduction. The 17 Sustainable Development Goals specified in the 2030 Agenda regard the travel guidelines on national scale that have less economic impact, generally chosen by users which deals with the environmental economic and social development. Specifically, it involves Goal 9, which concerns sustainable infrastructures and transportation, as well as Goal 13 about climate change [45–48].

A crucial contribution to this issue is provided by LARG approach [44–47]. The theme defined is to identify an approach that allows to support the choices to be made in the programming phase, and therefore which solution to adopt between several alternatives of railways, among others HSR, contributing to the achievement of the SDGs. It is useful to refer to a paradigm (one or more) that in its various articulations has been built over time. The approach was first developed in the automotive industrial sector and was then extended to other industrial sectors.

Within the paradigm LARG, four main components exist specified by the industry sector: *Lean***,** removal of the redundant elements; *Agile***,** ability to react to demand; *Resilience*, ability to respond to the risks; and *Green,* ability to concur positively in the environment.

This paper takes into account the component Green [44] it is possible to deepen the effects of the environmental impacts due to modal share, in terms of emission of greenhouse gases (GHG), especially regarding carbon dioxide equivalent (CO_{2eq}). The emissions assessment framework consists of distinct phases that categorize environmental impacts based on the energy life cycle, that are: Well-to-Tank (WTT), comprising energy

extraction, refining, and distribution; Tank-to-Wheels (TTW), involving the energy consumption and direct exhaust emissions during vehicle movement; and the broader Life Cycle Assessment (LCA) [49], which accounts for the manufacturing, maintenance, and disposal of both vehicles and infrastructure.

This study is focused on the Well-to-Tank (WTT) cycle as it regards the operational phases of the transport modes. This phase of analysis concerns the energy consumed for traction and the associated GHG emissions, enabling a direct comparison between the mode-service of HSR and aircraft.

The paper analyzes several scenarios of impacts caused by the choices of travelers' behavior, thus observed demand flows provided by Telco data, thus a current scenario. In addition, the impacts of modal share due to simulated scenarios are presented, obtained with a sensitivity analysis, hence on one hand from hypothesized improvements on the railway line and implementations of the services offered, therefore a share of demand diverted from the air mode. On the other hand, it is considered a decreasing of demand flows (from rail) towards air mode, hypothesized due to decreasing in the supply of service and/or infrastructural failures, which cause a share of demand regarding air mode. The results show that the air mode impacts with a quantity of GHG more 2.5 times than rail mode of transport.

The paper is organized as follows. Section 1 contains an analysis of the basic state of the art regarding the research topic of the paper itself. Section 2 concerns the specification of the methodology for GHG emission estimation from different transport modes in terms of CO_{2eq} emissions, by means of the indicator provided by the Agenda 2030, as reported in Goal 13, by using the EcoPassenger software provided by UIC [50]. Section 3 deals with the application of the methodology specified within Sect. 2. Therefore, the context of application was specified, namely the Italian Rome-Milan relationship, chosen because it results to be the national reference link in terms of flows and supplied services. The quantity of GHG emitted on the relationship has been calculated for different scenarios, firstly the actual scenario with Telco data. A sensitivity analysis is performed in order to estimate possible different future scenarios of the transport system among this relationship. The last section contains some final considerations.

This work is aimed at supporting transport planners and decision-makers in the evaluation of future investment in HSR lines and services, by means of methodological and modelling tools to assess the potential GHG saving and possible future investments in HSR infrastructure and services.

2 Methods to Estimate the Emissions

The chapter contains the specification of the methodology for calculating GHG emissions from different transport modes in terms of CO_{2eq} emissions, in addition it is included the method for calculating GHG emission parameters.

2.1 Model Specification

The methodology aimed at the calculation of total emissions of GHG and thus the consequent total emission saved due to the modal diversion of passengers from air to HSR mode, is presented in this section.

The data concerning the observed demand flows, for both modes of transport, for the actual scenario is Telco data. It is worth noting that the different scenarios are obtained with sensitivity analysis, reproducing possible variations in the infrastructural or services systems.

The procedure of calculation is carried out through the following equations. The demand flows, for each mode of transport (rail and air), are provided by Telco data (d_m), extracted from big data. The Eq. (1) in the paper is not specified and applied, because is applied the pivoting ±10%. Therefore, the analysis was executed considering 3 different scenarios:

1. Current scenario, C: in which the emission of CO2 (Eq. 2) is related to the demand flows observed;
2. Scenario rail, R: in which there is the hypothesis of diversion of 10% demand from air, towards rail (Eq. 2);
3. Scenario air, A: in which there is the hypothesis of diversion of 10% demand from rail, towards air (Eq. 2).

The calculation of the difference between rail and air transport modes, in terms of emitted CO_2, is estimated with the Eq. (3), thus is the difference between CO2 emitted by the scenario 1 or 2 and the actual scenario.

$$d_{m,w} = f(\mathbf{x}; b) \tag{1}$$

$$CO_{2\,m,w} = f\left(d_{m,w}; \mathbf{x}\right) \tag{2}$$

$$DCO_2 = f\left(CO_{2\,m,C};\ CO_{2\,m,w}\right) \tag{3}$$

With:

$d_{m,w}$ = demand flows referred to the mode *m* and the defined scenario w;
$\mathbf{x}$ = vector of attributes;
$CO_{2\,m,w}$ = carbon dioxide emissions of the mode *m* referred to the scenarios R and A;
$w \in \{C, R, A\}$, defined scenario.

2.2 Parameter Calculation Specification

All the components of greenhouse gas (GHG) produce a particular global warming effect.

Among the seventeen Sustainable Development Goals (SDGs) defined, HSR generates direct and indirect impacts on the following ones.

There exist direct impacts and indirect impacts generated by the transportation sector. On one hand, the direct impacts are related to the Goals 9 “Industry, Innovation and Infrastructure” and Goal 11 “Sustainable Cities and Communities”. On the other hand, the indirect impacts are related to Goals 7 “Affordable and Clean Energy”, Goal 8 “Decent work and Economic Growth”, Goal 10 “Reduce Inequalities” and 13 “Climate Action”.

This paper deepens Goal 13 “Climate Action”, which is a call for urgent action to cope with climate changes and its impacts. It pursues on the strengthening of resilience,

improving climate awareness, as well as integrating climate measures into national policies. Particularly, it is considered the Target 2 "Integrate climate change measures into national policies, strategies and planning", as wells as in particular the estimation of the Indicator 2 "Total greenhouse gas emissions per year". The latter indicator measures the total greenhouse gas emissions (GHG) per year of the actors which participates at the United Nations Framework Convention on Climate Change (UNFCCC) [51].

The total GHG emission, according to the document of United Nations Agenda 2030 [52, 53] is calculated by taking into account the emissions of direct GHGs:

- Carbon dioxide (CO_2)
- Methane (CH_4)
- Nitrous oxide (N_2O)
- Perfluorocarbons (PFC_x)
- Hydrofluorocarbons (HFC_x)
- Sulphur hexafluoride (SF_6)
- Nitrogen trifluoride (NF_3)

It is important to underline that carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O) are characterized by a high global warming potential.

The calculation of total GHG emissions is composed of two main steps (equations), established in the IPCC document: the mass of each component that emitted gas, and subsequently, a weighted summation using the respective GWPs is applied to meet the requirements of Indicator 13.2.2 Eq. 3.

$$E = \Sigma (G_i * GWP_i) \tag{4}$$

with:

E = GHG total emission;
G_i = Gas $mass_i$, total physical quantity emitted (in tonnes or megatonnes) of the *i* greenhouse gas (GHG) over a year, calculated by individual country for the various economical sectors (energy, transport, agriculture, waste, etc.);
GWP_i, Global Warming Potential specific to gas i, parametrated on the carbonical anhydride molecule. The UN specifies that this factor is useful to convert the impact of any gas into the same unit of measurement.

Indicator 13.2.2 of Agenda 2030 requires the calculation of total greenhouse gas (GHG) emissions by converting the different gases through their respective global warming potential (GWP), in full compliance with UNFCCC guidelines [51]. In this study, the same methodological principle was applied to quantify the carbon footprint of passenger mobility: the EcoPassenger model already aggregates within itself the different climate-altering factors associated with electricity production and fossil combustion. The unit parameters expressed in g/pkm and extracted from the model already include this conversion to equivalent CO_2 making the dynamic calculation based on Telco data fully consistent with the integrated approach required by Goal 13.

$$\text{Emissions} = \text{Activity Data (x)} * \text{Other Factor } (\beta) \tag{5}$$

With: Activity Data (x): quantitative attribute about activity, for instance tons of fuel burned, kilowatt-hours of electricity used, or kilograms of material consumed; Emission parameter (β): a parameter that converts activity data in quantity of total GHG emitted (e.g., per GJ of energy).

In order to pursue the calculation of the indicator specified by Agenda 2030, thus the indicator 13.2.2, it is used the tool provided by UIC and specified in the Methodology Report, available on the EcoPassenger website, for both components of the Eq. (4). The latter is a software aimed at the calculation of the GHG emitted in a specific context, to be specified.

2.3 Activity Data (X)

Passenger transportation in Europe is provided by different transport modes and services. Within the EcoPassenger software the most important modes of transport using ordinary vehicle types and propulsion systems are taken into account.

As elaborated in the EcoPassenger methodology (Chapter 1.2 [52]), the specified framework focuses uniquely on the emissions related to the operational phases (Well-to-Wheel), thus results to be the Activity data component of Eq. (2). It has been chosen to not include the environmental contributions of the manufacturing, maintenance, and disposal phases of vehicles, hence the construction and maintenance of transport infrastructures. Therefore, the framework does not consider the raw material extraction parameters included in the indicator calculation of Goal 13.

EcoPassenger does not consider other emission or resource consumption parameters because it is a model focused exclusively on operational energy and direct and indirect atmospheric emissions from travel, deliberately excluding the construction of means and infrastructure to ensure homogeneity and immediate data compatibility between the different modes. It considers CO_2 (and equivalent greenhouse effect), nitrogen oxides (NO_x), particulate matter (PM), and non-methane hydrocarbons (NMHC) because the latter are the direct products of combustion or mechanical wear (brakes/wheels) related to the travel itself (Table 1).

However, the methodology proposed by UIC software does not include each of the chemical elements specified in the UN report of Agenda 2030, Goal 13.2.2. The following table shows the elements included in the calculation of GHG by UIC software.

2.4 Other Factors (β)

The Emission parameter is provided by the EcoPassenger Methodology report, specified as “Radiative Forcing Index” - RFI factor, which takes into account the climate effects of the GHG emission component not included in the methodology (in particular nitrogen oxides, ozone, water, soot, sulphur), especially for emissions in high altitudes and dependent to the distance in km), and are specified in the following table. The result, including this parameter, allows to calculate the quantity of CO_2 which would have to be emitted and causing the identical global warming effect, as averaged globally, as the several pollutants collectively.

Considering the component of GHG emission, according to the document of United Nations Agenda 2030, in the following the elements not considered are described.

Table 1. GHG component included in EcoPasssenger (elaboration from: [50]).

GHG component	Description	Motivation for inclusion
CO_2	Carbon dioxide emissions	main indicator for GHG effect
NO_x	Nitrogen oxide emissions	Acidification, eutrophication, eco-toxicity, human toxicity, summer smog
PEC*	Primary energy consumption	Main indicator for resource consumption
NMHC	Non-methane hydro carbons	Human toxicity, summer smog
PM	Exhaust particulate matter from vehicles (combustion), mainly $PM_{2.5}$ and from energy production and provision. Composition: all particle sizes, about 80% $PM_{2.5}$ and PM_{10} (by mass)	Human toxicity, summer smog

PEC*: Well-to-Wheels energy consumption, displayed as "liter petrol equivalent" (1 L = 32.309 MJ.

Table 2. RFI factor individuation

Distance range (km)	RFI factor (average)
500	1.27
625	1.47
750	1.60
1,000	1.87
>1,000	2.50

Therefore, environmental components, for instance land use, noise pollution, and stratospheric ozone depletion were excluded from the framework of this work. Regarding electrified rail transit, the model did not include the radiological risks and waste management challenges associated with nuclear power generation. Furthermore, the current iteration does not take into account non-CO_2 GHG, specifically methane (CH_4) and nitrous oxide (N_2O), as CO_2 represents the main part of GHG emissions within the transport sector. While $CO_{2,eq}$ emissions considered in CH_4 and N_2O are just about 1% higher than pure CO_2 emissions for conventional fuels, this variance increases to a range of 2% to 10% for biofuels and electricity generation, Finally, particulate matter (PM) quantitative determination is rigorously confined to tailpipe and combustion emission exhaust, thus excluding non-exhaust emissions obtained from mechanical abrasion and suspension.

2.5 Query Set up in EcoPassenger for GHG Estimation

In the following it is specified the query to be set up in order to estimate the EF component (Eq. 2).
Step 1. Starting and arrival point: city, railway station, airport; date of departure or arrival.
Step 2. Car, Rail, Flight specifications
2.1. Car Specifications: Vehicle class (middle class/upper class); Engine: Gasoline/Diesel Euro 1–6, Diesel conventional, GPL Euro 1–6, Electric Vehicle. no. of passengers in the car: 1–5, European average utilization (1.5 passengers). The fuel consumption is specified in "Tab. 2–5" of the UIC report about EcoPassenger methodology.

2.2. Rail/Flight Specifications: Utilization per train and flight (Average load factor, Maximum load factor), global warming extra charge of flight (CO_2-emissions with climate factor, CO2-emissions without climate factor).

Concerning the typologies of trains, are considered the following in relation to the route, thus the service: electric and/or diesel for intercity and regional/urban services, electric for high speed services. The fuel and energy consumption is specified in "Tab. 2–7" of the UIC report about EcoPassenger methodology [50] On the other hand, for the flights (short and medium distances in Europe) are considered, by the software, the following typologies of airplanes: B737 family and A319/321 family, and takes into account the different phases of the flight. The latter element is crucial for shorter distances because the greatest consumption occurs in the take-off and landing operations of the flight. However, there are not considered the movements between airports for reasons of repositioning airplanes for following services. The fuel (kerosene) consumption is specified in Tab. 2–9 of the UIC report about EcoPassenger methodology [50].
Step 3. Feeder
3.1. Feeder to and from the airport (car only, train only, taxi);

3.2. Electricity Mix of Railways (national production mix, railway mix including Green Certificates).
Results. Calculated parameters:

- Carbon dioxide CO_2 (greenhouse-gas, global warming), which includes emissions associated with the production of electricity/fuel;
- Energy resource consumption PEC (resource consumption/primary energy), which includes emissions associated with the production and distribution of electricity/fuel;
- Particulate matter PM (human toxicity), related to the combustion (fuel), including production of fuel or electricity power;
- Nitrogen oxides NO_x (acidification, nutrification, summer smog/ozone), including production of fuel or electricity power;
- Nonmethane hydrocarbons NMHC (summer smog/ozone), including production of fuel or electricity power (Fig. 1).

It is important to underline that each of the other specification, concerning the Eco-Passenger GHG estimation, not reported in this work are clearly explained in Report of UIC about the methodology [50].

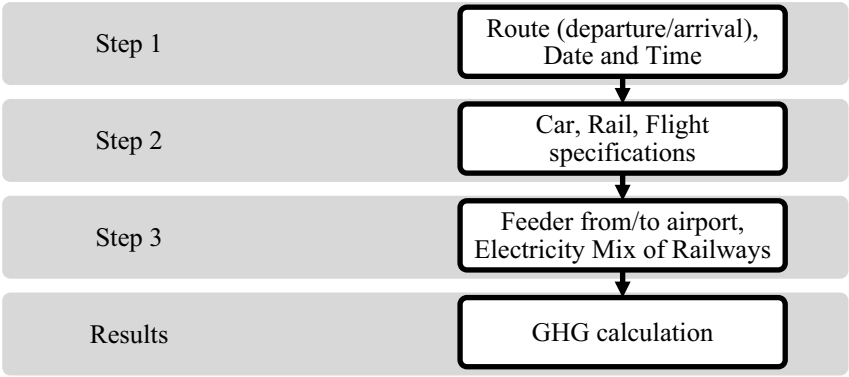


Fig. 1. Description of the specification phase in EcoPassenger

3 Case Study: Italian High Speed Rail

3.1 Experimental Test Site

The HSR line between Rome and Milan (Italy) was selected as it is the main Italian relationship in terms of volume of passengers and of services. The relationship includes the alternative of air services; however, this work does not investigate the mode choice level (Fig. 2).

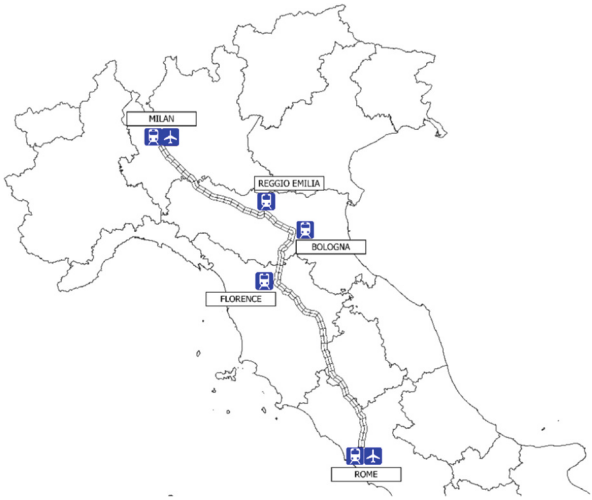


Fig. 2. Experimental test site: Rome-Milan link (Italy)

The case study concerns the Rome-Milan relationship (Italy). On one hand, on the railway side, has been opened in 2009. Consequently, it triggered both intra-modal and inter-modal competition, leading to a significant diversion of travel demand from air

and conventional rail services. Furthermore, induced travel demand emerged as a result of several factors, most notably the increased level of service offered by HSR (e.g., the reduction of in-vehicle travel time). Traffic data observed before and after the HSR opening reveals that passenger flows in 2008 were approximately 2.4 million/year for the air mode and 1.0 million/year for the conventional rail mode. By 2018, passenger flows shifted to 1.2 million/year by air and 3.6 million/year by HSR. Currently, two HSR companies operate along this corridor: Trenitalia, which supplies HSR services alongside conventional and local rail services, and Italo Nuovo Trasporto Viaggiatori (Italo NTV), which exclusively supplies HSR services. These two companies compete with each other at service and run choice levels, while jointly competing at the mode choice level with airline companies. As of 2024, the HSR services stands at approximately 45 runs/day (per direction) for Trenitalia and 32 runs/day for NTV. These scheduled runs are characterized by various operational elements, such as the typology of service (e.g., non-stop or shuttle runs with intermediate stops), in-vehicle travel time, and dynamic fare structures offered to travelers. Conversely, concerning the air mode, approximately 23 flights/day are operated by airline companies between the airports of Milan Linate and Rome Fiumicino.

3.2 Activity Data Estimation

The section contains the calculation of the emission parameters, by means of the EcoPassenger software, provided by UIC [50] (Table 3).

Table 3. Calculation of the emission parameters

Emission parameter	[UoM]	Train (HSR)	Car	Air
Carbon dioxide	[kg]	10.3	63.1	141.7
Energy resource consumption	[l.g.e.]	6.1	28.0	49.0
Particulate matter	[g]	2.2	2.0	10.5
Nitrogen oxides	[g]	25.8	345.8	565.4
Nonmethane hydrocarbons	[g]	1.4	20.5	48.9

With: UoM = unity of measure, kg = kilograms, g = grams l.g.e. = liter gasoline equivalent.

3.3 RFI Factor Individuation

As reported in Table 2 it indicates the RFI factor in relation to the distance of the experimental test site specified in Sect. 3.1, thus Rome-Milan relationship, and results to be 567 km for the railway as well as 472 km for the air link. Consequently, for both modal alternatives the RFI factor results to be 1.27.

3.4 Scenarios: Emission Estimation

The scenarios of the case study are calculated considering a sensitivity analysis with respect to the actual scenario, for both rail and air modes of transportation.

The actual scenario in terms of demand flows is provided by the results of an analysis for obtaining Telco data in the specified relationship, thus is 593.567 passengers for rail mode and 109.500 passengers for air mode. Table 4 contains the results of the GHG emissions estimation. Concerning the rail mode results to be 7.764,4 tons, whereas the GHG emissions for the air mode are 19.705,5 tons. The total GHG emissions for the actual scenario are 27.470,0 tons.

Regarding the scenarios are conducted by means of a sensitivity analysis, as specified in Sect. 2.

The scenario 1 is obtained by hypothesizing the diversion of the demand flows of about 10% from air towards rail, and results 604.517 passengers for rail mode and 98.550 passengers for air mode. Table 4 contains the results of the GHG emissions estimation. Concerning the rail mode results to be 7.907,7 tons, whereas the GHG emissions for the air mode are 17.735,0 tons. The total GHG emissions for scenario 1 are 25.642,6 tons.

The scenario 2 is obtained by hypothesizing the diversion of the demand flows of about 10% from rail towards air, and results 534.210 passengers for rail mode and 168.857 passengers for air mode. Table 4 contains the results of the GHG emissions estimation. Concerning the rail mode results to be 6.988,0 tons, whereas the GHG emissions for the air mode are 30.387,3 tons. The total GHG emissions for scenario 2 are 37.375,3 tons.

Table 4. GHG estimation for the different scenarios

Scenarios	Mode	Activity data	RFI factor	Demand flows	GHG emission	GHG emission
Actual		[kg]		[pax]	[kg]	[tons]
	Train	10,3	1,27	593.567	7.764.449,9	7.764,4
	Air	141,7	1,27	109.500	19.705.510,5	19.705,5
			Total	(actual)	27.469.960,4	27.470,0
1 (+10%)	Train	10,3	1,27	604.517	7.907.686,9	7.907,7
	Air	141,7	1,27	98.550	17.734.959,5	17.735,0
			Total	(1)	25.642.646,3	25.642,6
2 (-10%)	Train	10,3	1,27	534.210	6.988.004,9	6.988,0
	Air	141,7	1,27	168.857	30.387.282,9	30.387,3
			Total	(1)	37.375.287,8	37.375,3

4 Discussion and Conclusions

The paper presents an analysis of GHG production of two transport modes, such as air and rail, and a comparison between the actual and the simulated scenarios of GHG emissions in a case study of passenger mobility along the relationship Rome-Milan (Italy). The analysis' results show the relevant variations of the externalities under consideration

according to the specific sensitivity cases. In particular, the impact of the actual scenario, compared to the simulated ones, demonstrates the massive carbon footprint of the air mode despite its lower passenger flows, producing 19,705.5 tons of GHG against the 7,764.4 tons produced by the rail mode. Moreover, Scenario 1, which hypothesizes a 10% increase in passenger demand of Rail mode of transport due to service improvements, would lead to an overall emissions decreasing, reaching 17.735,0 tons for air and 7.907,7 tons for train. It is observed that scenario 1 is the best in terms of GHG saving and less emissions, whereas scenario 2 shows that the air mode impacts 4.4 times the rail mode of transport, in terms of GHG emissions.

It is clear that the diversion of passengers towards the HSR mode allows a relevant quantity of GHG savings compared to air mode, underling the environmental sustainability of rail transport mode. A crucial role is played by the travel demand; hence, it is important to correctly evaluate the HSR travel demand, because it supports the achievement of UN Sustainable Development Goal 13.

Acknowledgments. This study was carried out within the MOST – Sustainable Mobility National Re-search Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1033 17/06/2022, CN00000023). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

References

1. UIC: High Speed Traffic in the World – Passenger-km by Country (2025). https://uic.org/IMG/pdf/20250618_high-speed_traffic_passkm.pdf. Accessed 1 Feb 2025
2. European Commission website. https://commission.europa.eu/strategy-and-policy/recovery-plan-europe_en. Accessed 04 Apr 2026
3. IlSole24Ore webpage. https://lab24.ilsole24ore.com/cop27-dati-CO2-mondo/. Accessed 26 Feb 2024
4. Cascetta, E.: Transportation Systems Engineering: Theory and Methods, vol. 49. Springer, New York (2013). https://doi.org/10.1007/978-1-4757-6873-2
5. de Dios Ortúzar, J., Willumsen, L.G.: Modelling Transport. Wiley, Hoboken (2024)
6. Ben-Akiva, M.E., Lerman, S.R.: Discrete Choice Analysis: Theory and Application to Travel Demand, vol. 9. MIT Press (1985)
7. Cantarella, G.E., Watling, D., De Luca, S., Di Pace, R.: Dynamics and Stochasticity in Transportation Systems: Tools for Transportation Network Modelling. Elsevier, Amsterdam (2019)
8. Moschella, M.: Diodos 2: Approfondimenti sulla mobilità terrestre. Fondazione Filippo Caracciolo (2025). https://fondazionecaracciolo.aci.it//app/uploads/2025/01/Diodos_2_Approfondimenti-sulla-mobilita-terrestre.pdf
9. Musolino, G., Cartisano, A., Fortugno, G.: Methodologies for sustainable development of TEN-T/RFC corridors and core ports: estimation of time-series economic impact. In: Gervasi, O., et al. (eds.) ICCSA 2023. LNCS, vol. 14110, pp. 551–562. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-37123-3_38
10. Russo, F., Moschella, M.: Mode choice at national level: revealed and stated preference survey for high-speed rail realization. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15894, pp. 359–374. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97648-3_24

11. Russo, F., Musolino, G.: Methodologies for sustainable development of TEN-T/RFC corridors and core ports: economic impacts generated in port-related areas. In: Gervasi, O., et al. (eds.) ICCSA 2023. LNCS, vol. 14110, pp. 515–526. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-37123-3_35
12. Russo, F., Pellicanò, D.S.: Methodologies for sustainable development of TEN-T/RFC corridors and core ports: the role of governance in the export time optimization. In: Gervasi, O., et al. (eds.) ICCSA 2023. LNCS, vol. 14110, pp. 622–634. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-37123-3_43
13. Russo, F., Chilà, G., Zito, C.: Methodologies for sustainable development of TEN-T/RFC corridors and core ports: settlement capacity of industrial firms in port related areas. In: Gervasi, O., et al. (eds.) ICCSA 2023. LNCS, vol. 14110, pp. 539–550. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-37123-3_37
14. Cascetta, E., Nuzzolo, A., Biggiero, L., Russo, F.: Passenger and freight demand models for the Italian transportation system. In: Proceedings of the 7th World Conference on Transport Research, vol. 2 (1995)
15. Nuzzolo, A., Russo, F.: Departure time and path choice models for intercity transit assignment. In: Travel Behaviour Research: Updating the State of Play, Juan De Dios Ortuzar, David Hensher, Sergio Jara Diaz. Elsevier, Amsterdam (1998). ISBN 0 08 043360X
16. Russo, F., Rindone, C.: Regional transport plans: from direction role denied to common rules identified. Sustainability **13**(16), Article no. 9052 (2021). https://doi.org/10.3390/su13169052
17. Russo, F., Moschella, M., Musolino, G.: Railway demand evaluation: HSR induced component. In: Gervasi, O., Murgante, B., Garau, C., Taniar, D., Rocha, A.M.A.C., Faginas Lago, M.N. (eds.) ICCSA 2024. LNCS, vol. 14822, pp. 173–187. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-65318-6_12
18. Russo, F., Musolino, G., Moschella, M.: High-speed rail travel demand analysis: pilot experimentation on generation model. Transp. Res. Procedia **90**, 822–829 (2025)
19. Russo, F., Rindone, C., Maiolo, G.A.: High-speed railway planning for sustainable development: the role of length between conventional line and straight length. Future Transp. (2025)
20. Giunta, M.: Trends and challenges in railway sustainability: the state of the art regarding measures, strategies, and assessment tools. Sustainability **15**(24), 16632 (2023)
21. Giunta, M.: Data-driven track geometry defects localization and strategies for preventive maintenance: a case study. Transp. Res. Procedia **90**, 742–749 (2025)
22. Giunta, M., Leonardi, G.: Framework for life cycle railway sustainability assessment: a methodological approach based on advanced methods and tools. In: Gervasi, O., Murgante, B., Garau, C., Taniar, D., Rocha, A.M.A.C., Faginas Lago, M.N. (eds.) ICCSA 2024. LNCS, vol. 14822, pp. 233–244. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-65318-6_16
23. Panuccio, P.: High speed rail stations in territorial areas: sustainable development and quality of new landscape. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15894, pp. 434–451. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97648-3_29
24. Russo, A., Tesoriere, G., Rindone, C., Campisi, T.: Inter-urban transit connections: a network analysis in an Italian region. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15894, pp. 375–390. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97648-3_25
25. Di Gangi, M., Russo, F.: Potentiality of rail networks: integrated services on conventional and high-speed lines. Computers in Railways XVIII. WIT Trans. Built Environ. **213**, 101–112 (2022)
26. Di Gangi, M., Russo, F.: Design of hybrid rail services on conventional and high-speed lines. Int. J. Transp. Dev. Integr. **7**(2), 113–121 (2023)

27. Sassine, C.: Electric trains for reduction of fuel consumption and emissions from the freight transport in Jordan. PLoS ONE **20**(5), Article no. e0323121 (2025). https://doi.org/10.1371/journal.pone.0323121
28. Baran, J., Molková, T.: Connecting regional railway transport on line no. 2040 and future high-speed transport in the Vysočina Region using an algorithm. Transp. Res. Procedia **84**, 111–118 (2025). https://doi.org/10.1016/j.trpro.2025.04.117
29. Di Gangi, M., Belcore, O., Polimeni, A.: Rail transport across the Strait of Messina: some policies to improve the service. Open Transp. J. **18**(1), Article no. e26671212285520 (2024). https://doi.org/10.2174/0126671212285520240105073430
30. Dufek, J.: A case study of traffic volumes based on GIS distance analysis of shortest route data using own SHORTPATH toolbox. Transp. Res. Interdiscip. Perspect. **34**, Article no. 101744 (2025). https://doi.org/10.1016/j.trip.2025.101744
31. Švarcmajer, M., Ivanović, D., Rudec, T., Lukić, I.: Application of graph theory and variants of greedy graph coloring algorithms for optimization of distributed peer-to-peer blockchain networks. Technologies **13**(1), 33 (2025). https://doi.org/10.3390/technologies13010033
32. He, S., Mei, L., Wang, L.: The dynamic influence of high-speed rail on the spatial structure of economic networks and the underlying mechanisms in Northeastern China. ISPRS Int. J. Geo-Inf. **10**(11), Article no. 776 (2021). https://doi.org/10.3390/ijgi10110776
33. Musolino, G., Cartisano, A., Fortugno, G.: Structural factors for a third-generation port: planning interventions for mechanical logistics in Gioia Tauro, Italy. Urban Marit. Transp. XXVII **204**, 55–66 (2021). https://doi.org/10.2495/UT210051
34. Russo, F., Musolino, G., Assumma, V.: An integrated procedure to estimate demand flows of maritime container transport at international scale. Int. J. Shipp. Transp. Logist. **6**(2), 112–132 (2014). https://doi.org/10.1504/IJSTL.2014.059566
35. Giunta, M., Barrile, V., Leonardi, G., Genovese, E.: Comprehensive railway track monitoring using unmanned aerial systems (UASs) and building information modeling (BIM). In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15894, pp. 407–419. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97648-3_27
36. Giunta, M., Frontera, P., Rouisse, M.E.: Circular economy in rail track: estimating the environmental burdens of the use of recycled material and EoL processing through life cycle assessment. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15894, pp. 391–406. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97648-3_26
37. Pellicanò, D.S., Trecozzi, M.R.: Methodologies for sustainable development of TEN-T/RFC corridors and core ports: public incentives for industrial activities location in port related areas. In: Gervasi, O., et al. (eds.) ICCSA 2023. LNCS, vol. 14110, pp. 635–646. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-37123-3_44
38. Praticò, F.G., Fedele, R.: Economic sustainability of high-speed and high-capacity railways. Sustainability **15**(1), 725 (2023)
39. Rindone, C., Russo, A.: A network analysis for HSR services in the South of Italy. In: Gervasi, O., Murgante, B., Garau, C., Taniar, D., C. Rocha, A.M.A., Faginas Lago, M.N. (eds.) ICCSA 2024. LNCS, vol. 14822, pp. 217–232. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-65318-6_15
40. Rindone, C., Russo, A.: Centrality analysis for railway network planning: constrained measures for high speed. Transp. Eng. **22**, 100393 (2025). https://doi.org/10.1016/j.treng.2025.100393
41. Russo, F., Sgro, D., Musolino, G.: Estimating run choice models with innovative data collection: day-to-day tickets evolution in high speed rail. Transp. Policy (2025). https://doi.org/10.1016/j.tranpol.2025.103909
42. Russo, F., Sgro, D., Musolino, G.: Sustainable development of railway corridors: methods and models for high speed rail (HSR) demand analysis. In: Gervasi, O., et al. (eds.) ICCSA

2023. LNCS, vol. 14110, pp. 527–538. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-37123-3_36
43. Russo, F., Sgro, D., Musolino, G.: Dynamic structure of fares for high speed rail services. In: Gervasi, O., Murgante, B., Garau, C., Taniar, D., Rocha, A.M.A.C., Faginas Lago, M.N. (eds.) ICCSA 2024. LNCS, vol. 14822, pp. 188–201. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-65318-6_13
44. Russo, F.: Which high-speed rail? LARG approach between plan and design. Future Transp. **1**(2), 202–226 (2021)
45. Prakongwittaya, S., Liangrokapart, J.: Integrating LARG measures to improve supply chain transparency performance. Asia Pac. Manag. Rev. **30**(1), 100324 (2025). https://doi.org/10.1016/j.apmrv.2024.09.002
46. Surmařová, S., Neumannová, M., Pařil, V., Vrána, M., Chmelík, J.: Do I really like to shift to rail? Influence of rail modernisation on passenger preferences. Res. Transp. Bus. Manag. **59**, 101284 (2025). https://doi.org/10.1016/j.rtbm.2024.101284
47. Li, D., Jiang, R., Lu, Z., Sun, S., Wang, L.: Does the construction of high-speed rail change the development of regional finance? Sustainability **15**(13), Article no. 10641 (2023). https://doi.org/10.3390/su151310641
48. UN—United Nations. Transforming Our World: The 2030 Agenda for Sustainable Development (2015). https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf. Accessed 24 July 2023
49. Lu, Y., Zhang, T., Qiu, S., Liu, X., Yu, X., Yao, H.: Carbon emissions reduction potentiality for railroad transportation based on life cycle assessment. High-Speed Railw. **1**(3), 195–203 (2023). https://doi.org/10.1016/j.hspr.2023.08.004
50. International Union of Railways: EcoPassenger methodology report. UIC (2026). https://uic.hafas.cloud/bin/help.exe/en?L=vs_uic&tpl=methodology
51. UN—United Nations. Global Indicator Framework for the Sustainable Development Goals and Targets of the2030 Agenda for Sustainable Development (2018). https://unstats.un.org/sdgs/indicators/Global%20Indicator%20Framework%20after%202023%20refinement_Eng.pdf. Accessed 24 July 2023
52. United Nations: Our Common Future. Report of the World Commission on Environmental and Development (1987)
53. United Nations Framework Convention on Climate Change: United Nations Framework Convention on Climate Change: Text. UNFCCC (1992). https://unfccc.int/resource/docs/convkp/conveng.pdf

An Assessment of Road Network Topology and Its Influence on Port Accessibility in Romania

Aura Rusca, Florin Rusca(✉), Alina Roman, Oana Dinu, Stefan Burciu, Iulia Dorobantu, and Denis Codroiu

National University of Science and Technology Politehnica Bucharest,
313 Spl Independentei, District 6, Bucharest, Romania
florin.rusca@upb.ro

Abstract. Maritime ports represent the primary gateway between a country's land-based transport system and the global marketplace. Their capacity and configuration in relation to freight flow demands constitute a central topic of research within the field. However, vulnerabilities within the terrestrial transport network expose these ports to disruptive factors of both internal and external origin. In the case of Romanian maritime ports, the bridges crossing the Danube ensure their connection to the national road network, but they may also, in the event of external interference, lead to the isolation of these ports and a consequent reduction in the volume of goods passing through port terminals. The application of a macro-level network simulation model allows for the assessment of the impact generated by the unavailability of any of these critical road network components. The results obtained are useful for decision-makers in developing transport network strategies that ensure redundancy, with beneficial effects that can be economically quantified.

Keywords: Accessibility · Land transport network · Macroscopic simulation · Network Topology

1 Introduction

The relationship between maritime ports and the land transport network within their hinterland is a critically important aspect in the development strategy of any country. This is also influenced by the role of maritime ports within supply chain structures. Therefore, beyond the handling and throughput capacity of ports, their performance is also affected by the performance of the land transport networks that connect them to their hinterland. As freight volumes continue to grow and supply chains become more integrated, the efficiency of land network has become a key concern for researchers and practitioners also.

The structure and main infrastructure elements of road networks play a critical role in ensuring an acceptable level of accessibility and network connectivity. Engineering structures such as bridges, tunnels, and viaducts often become sets of critical links, especially where natural barriers limit the availability of alternative routes. In extreme

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 585–597, 2027.
https://doi.org/10.1007/978-3-032-30530-5_37

cases, the failure of a single infrastructure component may lead to the partial isolation of key economic regions or transport hubs, as is the case with maritime ports.

The Romanian Road transport network provides a relevant example of this type of vulnerability. The Danube River, while representing an important transport corridor, also constitutes a major geographical barrier for road transport. Internal road crossings over the Danube are relatively few (only three at present) and are distributed along a stretch of more than 200 km, making them essential links between the maritime ports on the Black Sea coast and the rest of the national road network.

The occurrence of disruptions affecting the proper functioning of Danube bridges can alter freight routing patterns and increase travel distances and times for vehicles serving maritime ports. These increases are quantified by transport operators in monetary cost terms. In more severe scenarios, with significant impacts on the network, the reduced availability of these crossings leads to a limitation in the ability of ports to be efficiently connected to their hinterland. Despite these risks, the interaction between port performance and the vulnerability of road infrastructure has not been extensively explored, particularly at the macro-network level and in the context of Eastern European transport systems.

Existing research in the literature provides a range of results regarding port efficiency, hinterland connectivity, and transport network resilience. However, these topics are often addressed separately, without explicitly highlighting the interdependencies between them. Studies analyzing how disruptions to critical road infrastructure affect maritime ports are relatively limited, and while macro-level simulation approaches capable of capturing system-wide impacts do exist, they have not been specifically applied to the case of the Romanian road transport network. In response to these gaps, in this paper the authors investigate the vulnerability of Romanian maritime ports to disruptions affecting critical components of the road network. A macro-level transport network simulation model is used to assess the impact of the unavailability of road bridges across the Danube on freight flows directed toward maritime port terminals. The study aims, based on the conducted research, to provide a clear picture of the structural robustness of the existing road transport network in Romania.

The results of the research conducted by the authors contribute to the advancement of the transport field, particularly in relation to aspects such as the vulnerability and robustness of transport networks, highlighting the central role of critical infrastructure in maintaining port connectivity. From a practical perspective, the findings are useful for the planning of strategic transport infrastructure projects, as well as for informing public policy decisions aimed at increasing network robustness and reducing the potential impact of disruptive events on the performance of maritime ports.

Against this background, the study looks at disruptions in the transport network. The focus is on critical infrastructure elements. The analysis is carried out at a national scale. Special attention is given to maritime port accessibility. The research question is clear. How does the road network affect access to ports? Especially when disruptions occur. A national-level transport model is used. It helps track how traffic shifts to alternative routes. It also shows the extra costs for users. The contribution is mainly practical. It clarifies how vulnerability affects accessibility. It also supports better planning and resilience decisions.

The research is structured into six chapters. The first chapter outlines the context in which the study was conducted. Section 2, Literature Review, provides a critical analysis of the relevant bibliographic sources considered by the authors and identifies the existing research gap addressed by this study. The third chapter presents current methods for evaluating accessibility in territorial transport networks. Section 4, Modelling the Romanian Road Network, introduces the conceptual framework applied in modelling the national road network. In Sect. 5, a case study is developed using the PTV Visum platform, focusing on the influence of the operation of major Danube ports on the accessibility of Romanian maritime ports. Finally, Sect. 6, Conclusions, summarizes the main findings and outlines directions for future research.

2 Literature Review

The assessment of transport network exposure is based on the concepts of reliability and vulnerability of transport networks. Borrowed from the field of technical component reliability, these concepts allow the evaluation of the impact that disruptive events may have on the functioning of the transport network [1]. An important body of literature has focused on the vulnerability of transport networks and the consequences of disruptions affecting critical infrastructure elements. Berdica (2002) provided one of the earliest comprehensive discussions of road network vulnerability, emphasizing that failures of specific links may significantly affect network performance. Later studies proposed analytical approaches for identifying critical infrastructure components within transport systems [2]. For example, Jenelius, Petersen and Mattsson (2006) developed a framework for evaluating the exposure and importance of road network links [3], while Murray-Tuite and Mahmassani (2004) examined methods for identifying vulnerable links whose disruption may strongly affect network accessibility [4]. Similarly, Taylor and D'Este (2007) proposed diagnostic approaches for detecting critical locations within transport infrastructure systems [5]. It is also worth noting the analysis of the influence that the unavailability of certain critical links has on network reliability. In this case, the quantification of the results is based on travel time as the main parameter [6].

The interaction between maritime or inland ports and their hinterland has been extensively analyzed in transport geography and logistics research [7–11]. Ports function not only as nodes in global maritime networks but also as gateways connecting maritime flows with inland transport systems. This interaction is assessed starting from the port gate, by evaluating how the transport network is connected to the port's internal infrastructure [12]. The ports efficiency and long-term development therefore depend on the performance of the terrestrial infrastructure that supports freight movements beyond the port area. As highlighted by Rodrigue, Comtois and Slack (2020), modern ports operate as integrated elements of complex logistics systems where road, rail and inland waterways ensure the continuity of supply chains between maritime and inland regions [14].

A significant contribution to the understanding of port–hinterland dynamics was made by Notteboom and Rodrigue (2005), who introduced the concept of port regionalization. Their work explains how ports evolve from simple transshipment points into integrated logistics platforms connected to regional transport corridors and inland terminals [7]. In this framework, the quality and reliability of inland transport infrastructure

become essential factors shaping the size and accessibility of port hinterlands (Notteboom, 2008) [8]. A complementary perspective is offered by Balletto, Borruso and Campisi (2022), who underline that port–hinterland relations are increasingly shaped not only by waterfront functions, but also by inner harbor areas and peripheral logistics spaces, where intermodal and warehousing activities support the inland extension of port operations [15]. Other studies consider statistical parameters related both to the transport network and to the economic dimension in order to evaluate the efficiency of ports in relation to their served hinterland [10].

Although previous studies have examined both port–hinterland relations and the vulnerability of transport networks, relatively limited attention has been paid to the exposure of maritime ports to disruptions occurring in specific inland infrastructure elements. As situations may arise in which ports are connected to their hinterland through a limited number of transport links, they become particularly vulnerable when transport infrastructure becomes unavailable. This aspect is also relevant for Romanian maritime ports, whose road accessibility depends significantly on the bridges crossing the Danube River. The link between transport networks and ports has been widely studied. However, network vulnerability is less explored. This is especially true in relation to port accessibility. Many studies rely on economic indicators. Others focus on connectivity measures. Only a few consider disruption effects. This study builds on that gap. The research question is stated clearly. It focuses on the role of critical network elements. The contribution has two sides. First, it moves beyond static analysis. It introduces disruption scenarios. Second, it uses a modelling approach. This helps capture traffic redistribution. It also shows how accessibility is affected. In this way, the gap in the literature becomes clearer. The position of this research is also better defined.

3 Evaluating the Accessibility

The assessment of accessibility in transport networks represents one of the research directions in the field that has generated sustained interest among specialists. The first mathematical models aimed to quantify the potential of a node or an area to interact with other locations through the transport infrastructure, depending on travel cost and the spatial distribution of activities. An early definition described accessibility as the potential of opportunities that can be reached from a given location, mathematically formulated through a spatial interaction–type relationship [16]. The gravitational accessibility indicator can be expressed by the following relation:

$$A_i = \sum_{j=1} O_j f(c_{ij}) \tag{1}$$

where A_i is accessibility of node or zone i, O_j is magnitude of opportunities in zone j (population, jobs, freight volume, etc.), c_{ij} is generalized travel cost between i and j and finally $f(c_{ij})$ is impedance function reflecting the negative effect of distance or travel time. A derived form of the accessibility assessment model is represented by the cumulative opportunity indicator. Accessibility is defined as the number of opportunities that can be reached within a given cost or time threshold:

$$A_i = \sum_{j=1} O_j I(c_{ij} < C) \tag{2}$$

where $I()$ is an indicator function depending by the maximum acceptable cost or time threshold [17].

In modern transport modelling applications, accessibility is calculated using the generalized cost matrix obtained through shortest-path algorithms, integrated into network models implemented in specialized transport planning software. The application of macro-simulation models enables the estimation of traffic flows distributed across the transport infrastructure network. Within this framework, travel times can be determined as a function of the capacity of each individual network link, the allocated traffic volumes, and the topological and structural characteristics of the respective infrastructure elements. Building on this modelling framework, network accessibility can be evaluated using the outputs generated by macro-simulation models. These models produce origin–destination traffic flows and associated travel times across the transport network while accounting for factors such as congestion, link capacity and route choice behavior. As a result, the generalized travel cost between zones can be derived directly from the simulated network conditions.

Using these outputs, several accessibility indicators can be calculated. Gravity-based indicators rely on the generalized cost matrix obtained from the simulation, while cumulative opportunity measures identify the number of opportunities that can be reached within a given travel time threshold. Since travel times are determined based on traffic volumes obtained from the traffic assignment process, the results have a higher level of credibility, reflecting the actual operating conditions of the network. This aspect makes macro-simulation a useful tool for analyzing how changes or disruptions in transport infrastructure operation influence accessibility patterns across the network.

Building on the previous approach, accessibility can also be represented graphically through isochrone maps derived from macro-simulation outputs. An isochrone defines the area that can be reached from a given origin within a specific travel time threshold, using the travel times calculated on the simulated network. At the zonal level, isochrones are constructed from the generalized travel cost matrix obtained in the macro-simulation model. Isochrones represent spatial surfaces that group together locations within a network characterized by the same temporal distance from a reference point. In transport network analysis, specialized software tools enable the selection of one or several network elements (e.g., nodes, stations, or traffic zones), after which the travel time between these elements and all other relevant components of the network is computed, typically using shortest-path algorithms Based on the obtained travel time values, a set of isochrones can be graphically represented using these temporal attributes of the network, thereby providing a visual representation of the spatial distribution of accessibility of maritime ports in relation to the national territory. The territory for which the travel time from the considered origin is below a predefined threshold (e.g., 30, 60, or 90 min at the local level and 3, 6, and 9 h at the regional and national levels) is represented on the map as accessibility areas. Such graphical representations provide an intuitive way to visualize variations in accessibility across the transport network and to identify areas with limited connectivity or affected by changes in the operating conditions of transport infrastructure.

4 Modelling Romanian Road Network

From a topological perspective, the Romanian national road network is structured around a radial configuration centered on the Romanian capital, Bucharest. The national roads have the origin in the country's main city and extend toward the major regions, forming the primary transport corridors. These corridors are complemented by transversal connections that link adjacent areas, thereby contributing to the overall cohesion and connectivity of the network (Fig. 1).

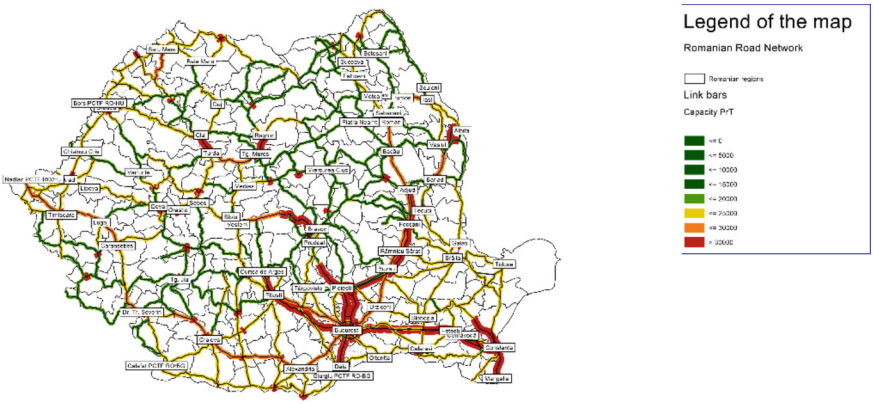


Fig. 1. Romanian road network (source: adapted and updated from [18]).

The motorway network further supports these connections by offering higher-capacity links between inland regions and the network's external gateways. In this context, the role of the road network in linking inland areas to the maritime ports along the Black Sea coast becomes clearly visible. By relying on a mix of routes with different capacity levels, freight flows generated in the hinterland are able to reach port terminals, facilitating the integration of maritime transport nodes into the broader national transport system. To model the accessibility of maritime ports, the road network is represented within a macroscopic model implemented in the PTV Visum software environment. The road capacity and speed limits are incorporated in accordance with the parameters used in the national transport models or in previous research [19–21].

The algorithm used for the assessment of road network topology and its influence on port accessibility in Romania is presented in Fig. 2. Using the traffic flow equilibrium assignment method, an initial set of traffic distribution values across the transport network is obtained. The equilibrium assignment method is used in macrosimulation models to distribute vehicle flows across routes within a transport network [22]. This method is a tool within the PTV Visum software and is based on Wardrop's first principle: "Every road user selects their route in such a way that the impedance on all alternative routes is equal, and switching to a different route would increase their personal travel time (user optimum).". Impedance is a modeling parameter that represents a proxy for the travel cost perceived by users. These variations are incorporated by software tool PTV VISUM into the model calibration and control parameters using the default values (the N

maximum number of steps is set to 20 and the maximum relative gap is set to 0.001). In the specific case of this research, given the national-scale nature of the model, impedance was expressed as travel time. The value of impedance varies dynamically with traffic volumes through volume–delay relationships. For the input data are used data from the last General Transport Master Plan of Romania [18].

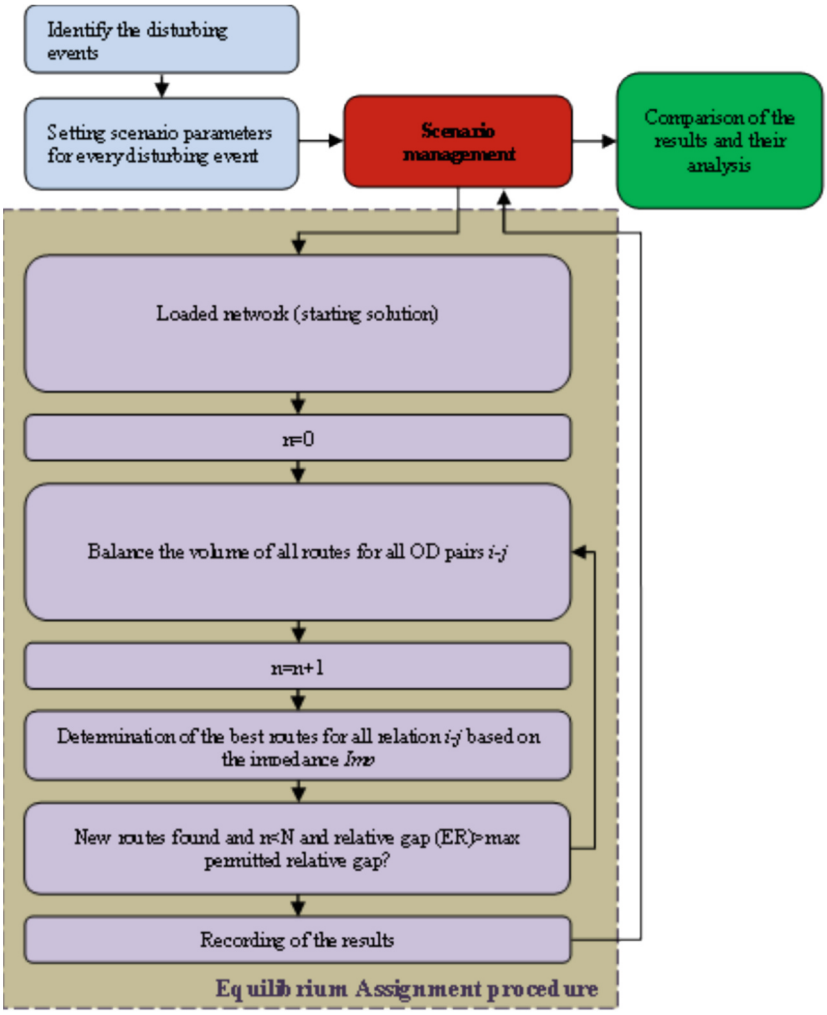


Fig. 2. The algorithm for evaluating the influence of disturbing events on port accessibility (source: adapted from [22]).

Based on these results, events characterized either by a high probability of occurrence or by a potentially significant impact on network performance are identified. Subsequently, a series of plausible scenarios are developed to represent such disruptions.

For each scenario, the virtual representation of the transport network is modified accordingly, and the traffic assignment procedure is repeated in order to simulate the redistribution of traffic flows under the new network conditions. The resulting traffic patterns are then evaluated through the construction of isochrone maps, which allow for a comparative assessment of the changes in maritime port accessibility across the different simulated scenarios.

5 The Study Case. The Influence of Danube Briges Functionality

Romanian maritime ports are connected to their hinterland through the national road network. From a spatial perspective, these ports are located in the south-eastern part of Romania, within the historical region of Dobrogea. Connectivity between this region and the rest of the country is ensured primarily through crossings over the Danube River. Due to the challenging topography, historical development patterns, and the considerable width of the Danube in the sector adjacent to the maritime ports, only three road bridges have been developed (Fig. 3).

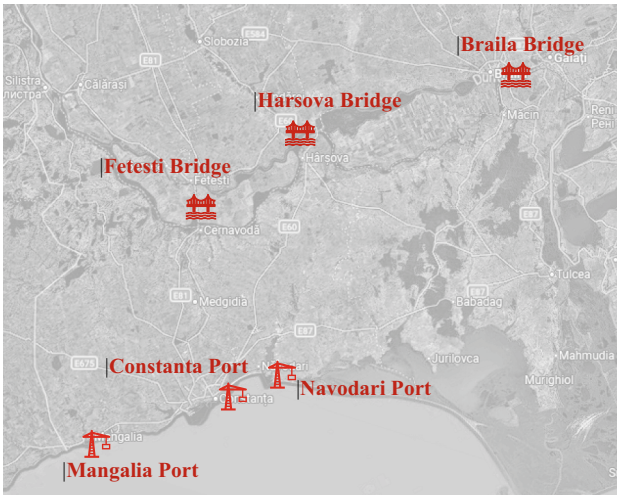


Fig. 3. Romania Maritime ports and Danube bridges

The operational performance of these bridges plays a critical role in determining the level of accessibility of Romanian maritime ports. Disruptive events—such as traffic accidents, maintenance works, or, in extreme cases, deliberate attacks—may lead to the partial or complete loss of functionality of one or more crossings. Although the resulting impact is inherently difficult to quantify, it can be systematically assessed through the development of scenario-based analyses reflecting different bridge availability conditions. The application of macroscopic traffic simulation models enables the evaluation of these scenarios, facilitating both quantitative and qualitative assessments of their effects on network performance and port accessibility.

The scenarios description are presented in Table 1. The analyzed three bridges are encoded as B1 for the Feteşti Bridge, B2 for the Hârşova Bridge, and B3 for the Brăila Bridge. The scenario numbering system employs a three-digit code reflecting the operational status of each bridge. A value of 1 indicates normal operation, while a value of 0 denotes the unavailability of the bridge due to a disruptive event.

Table 1. The macrosimulation scenarios

Scenario	Scenario description
S111	All three bridges are operational
S011	Bridge B1 is out of service, while the other two bridges remain operational
S101	Bridge B2 is out of service, while the other two bridges remain operational
S110	Bridge B2 is out of service, while the other two bridges remain operational
S100	Bridges B2 and B3 are out of service, while the remaining bridge is operational
S010	Bridges B1 and B3 are out of service, while the remaining bridge is operational
S001	Bridges B1 and B2 are out of service, while the remaining bridge is operational

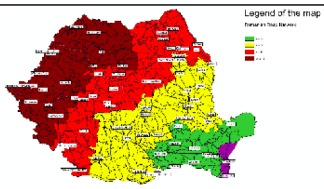
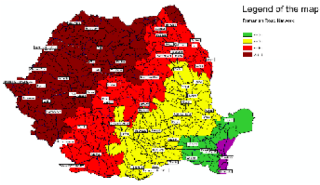
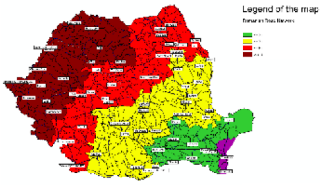
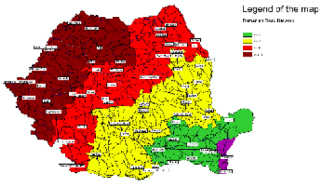
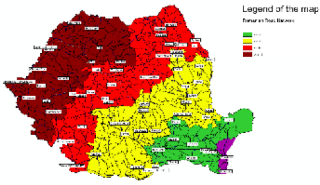
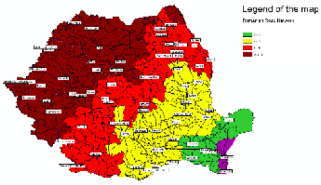
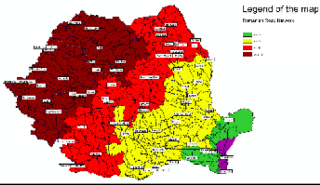
The results obtained with macrosimulation model are presented in Table 2. For every scenario is presented the accessibility map for time thresholds of 3, 6, and 9 h. Additionally, the percentage of territorial coverage is presented for the previously defined time thresholds.

The numerical results highlight that the unavailability of bridges B1, B2, or B3 leads to a direct degradation of maritime port accessibility, with the magnitude of the effect depending on the topological role of each bridge within the network. For example, the removal of bridge B1 results in a significant reduction in territorial coverage for the 3-hour threshold, indicating a rapid contraction of the ports' immediate hinterland. In the case where bridge B1 and B2 become simultaneously unavailable, the affected territory is increased, reflecting a major loss in port accessibility and the use of longer transport routes. Even in the case of disruptions affecting bridge B3, a negative impact is observed for the 6- and 9-hour analysis thresholds, confirming that each of these structures plays a critical role in maintaining efficient connectivity between the hinterland and maritime ports.

The analysis of the results highlights the critical importance of the Feteşti Bridge, coded as B1. In the event of a disruptive incident, the bridges included in the model—Hârşova Bridge and Brăila Bridge—play a redistribution role by absorbing part of the traffic originally assigned to B1, albeit with increased costs perceived by users.

Mitigation measures in this context include ensuring proper maintenance operations for all three bridges, encouraging users to also utilize network alternatives through neighboring Schengen countries (e.g., Bulgaria), and providing timely information regarding the operational status of these critical network elements.

Table 2. The results

Scenario	Macrosimulation output	Territorial coverage [km^2]			
		<3 h	<6 h	<9 h	>9 h
S111		31988	74296	67507	64679
S011		20169	58614	71921	87766
S101		31988	72102	69701	64679
S110		31988	72102	69701	64679
S100		31988	72102	69701	64679
S010		20169	58614	71921	87766
S001		17068	64311	75393	81698

6 Conclusions

The analysis of maritime port accessibility is of significant interest due to its economic implications. Maritime ports represent a key component of international supply chains. However, the performance of these transport nodes is not determined solely by the processes occurring within port areas. The connection between ports and their hinterland also plays a crucial role. The land transport network can either facilitate or hinder the connection between the port location and the areas from which freight flows originate or to which they are destined. In the case of road networks, it is essential to identify those critical infrastructure elements (such as bridges, tunnels, and viaducts) whose unavailability may lead to the isolation of network nodes or to the use of longer transport routes associated with higher costs. The use of macro-simulation models, combined with the representation of isochrones within the analyzed territory, constitutes a valuable analytical tool. This approach enables the assessment of the impact of disruptive events on the network, providing an integrated view of accessibility variations. It also facilitates the identification of critical points and vulnerable areas, offering a reliable basis for strategic decision-making.

In this study, the research focuses on the specific case of the Romanian road network that ensures the connection between maritime ports and their hinterland. Particular attention is given to the role of the Danube River, which effectively divides the network into two distinct regions connected by only three crossings considered critical. A set of scenarios was developed to evaluate the impact of the unavailability of each crossing individually, as well as in combined scenarios involving the simultaneous disruption of two links.

The obtained results allowed for the assessment of the relative importance of each crossing. The analysis highlights one particular link (the Feteşti Bridge) as having a major influence, while the impact of the unavailability of the other two bridges is less pronounced. This outcome is also supported by the topological structure of the network, which exhibits an egocentric configuration centered around the capital city of Romania.

The main limitations of the developed model are related to the nature of the data input. Traffic flow values were considered at an annual average level, while daily and seasonal variations were not included in the analysis. Future research aims to incorporate these temporal variations into both the modeling framework and the scenario analysis.

Overall, the proposed model represents a useful tool for the analysis of road transport networks and can support the justification of national development strategies by providing an increased level of credibility in the decision-making process.

Acknowledgments. The research work contained in this paper was supported within the frame of the project Innovative platform for the provision and management of public services in smart cities - SMART HUB, Code SMIS 120419.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Rusca, F., Roșca, E., Rusca, A., Roșca, M., Burciu, S.: The influence of transport network vulnerability for maritime ports. In: IOP Conference Series: Materials Science and Engineering, vol. 95, p. 012080 (2015)
2. Berdica, K.: An introduction to road vulnerability: what has been done, is done and should be done. Transp. Policy **9**(2), 117–127 (2002)
3. Jenelius, E., Petersen, T., Mattsson, L.G.: Importance and exposure in road network vulnerability analysis. Transp. Res. Part A Policy Pract. **40**(7), 537–560 (2006)
4. Murray-Tuite, P., Mahmassani, H.: Methodology for determining vulnerable links in a transportation network. Transp. Res. Rec. **1882**, 88–96 (2004)
5. Taylor, M.A.P., D'Este, G.M.: Transport network vulnerability: a method for diagnosis of critical locations in transport infrastructure systems. Transp. Policy **14**(2), 94–101 (2007)
6. Popa, M., Rosca, E., Rusca, F.V.: On transport network reliability. Transp. Probl. **7**(3), 127–134 (2012)
7. Notteboom, T., Rodrigue, J.P.: Port regionalization: towards a new phase in port development. Marit. Policy Manag. **32**(3), 297–313 (2005)
8. Notteboom, T.: The relationship between seaports and the intermodal hinterland in light of global supply chains. Discussion Paper No. 2008-10. OECD/International Transport Forum, Paris (2008)
9. Dinu, O., Dragu, V., Rusca, F., Ilie, A., Oprea, C.: Intermodal transport and distribution patterns in ports relationship to hinterland. In: IOP Conference Series: Materials Science and Engineering, vol. 227, p. 012038 (2017)
10. Roșca, E., Costea, I., Ilie, A., Petrović, M., Rusca, F.: A data envelopment analysis of inland ports' efficiency: evidence from the Romanian Danube ports. Systems **14**(2), 201 (2026)
11. Ducruet, C., Notteboom, T.: The worldwide maritime network of container shipping: spatial structure and regional dynamics. Prog. Hum. Geogr. **36**(3), 370–389 (2012)
12. Rusca, A., Dinu, O., Rusca, F., Roșca, E., Roșca, M., Ilie, A.: Evaluating the impact of using smart equipment for preselecting cargo vehicles in the maritime port access area. In: Santos, T.A., Soares, C.G. (eds.) Advances in Maritime Technology and Engineering (MARTECH 2024), pp. 81–86. CRC Press (2024)
13. Ducruet, C.: The geography of maritime networks: a critical review. J. Transp. Geogr. **88**, 102824 (2020). https://doi.org/10.1016/j.jtrangeo.2020.102824
14. Rodrigue, J.P., Comtois, C., Slack, B.: The Geography of Transport Systems, 5th edn. Routledge, New York (2020)
15. Balletto, G., Borruso, G., Campisi, T.: Not only waterfront. the port-city relations between peripheries and inner harbors. In: Gervasi, O., Murgante, B., Misra, S., Rocha, A.M.A.C., Garau, C. (eds.) ICCSA 2022. LNCS, vol. 13381, pp. 196–208. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-10548-7_15
16. Hansen, W.G.: How accessibility shapes land use. J. Am. Inst. Plan. **25**(2), 73–76 (1959)
17. Geurs, K.T., van Wee, B.: Accessibility evaluation of land-use and transport strategies: review and research directions. J. Transp. Geogr. **12**(2), 127–140 (2004)
18. Government of Romania, Ministry of Transport: General Transport Master Plan for Romania. Revised Final Version of the Report for Short-, Medium-, and Long-Term Planning. Bucharest (2015). https://mt.gov.ro/web14/strategii/master-plan-general-transport. Accessed 5 Mar 2026
19. Rusca, F.V., Rusca, A., Rosca, E., Oprea, C., Dinu, O., Ilie, A.: The East-European maritime ports hinterland and the influence of road transport network. In: Trends in Maritime Technology and Engineering (MARTECH 2022), vol. 2, pp. 85–90. CRC Press/Balkema, London (2022)

20. Dinu, O., Rosca, E., Rusca, F.: Inland waterway ports characteristics relevant for classification methodology purpose. In: Proceedings of the 3rd International Conference of Maritime Technology and Engineering (MARTECH), Lisbon, Portugal, pp. 53–59 (2016)
21. Rusca, F., Rusca, A., Rosca, E., Coman, C., Burciu, S., Oprea, C.: Machines **10**(10), 937 (2022). https://doi.org/10.3390/machines10100937
22. PTV Visum 2026: User Manual. PTV Group, Karlsruhe, pp. 396–401 (2026)

A LEAN Approach for High-Speed Railways: General Elements for Extending the Italian Network in the South

Corrado Rindone[1](✉) and Antonio Russo[2]

[1] DIIES, Università Mediterranea di Reggio Calabria, 89100 Reggio Calabria, Italy
corrado.rindone@unirc.it

[2] Dipartimento di Ingegneria e Architettura, Università degli Studi di Enna "Kore", 94100 Enna, Italy
antonio.russo@unikore.it

Abstract. High-Speed rail (HSR) infrastructures are developing all over the world. This represents a sustainable travel option according to sustainability goals. Recent works have proposed adopting the LARG (Lean, Agile, Resilient, Green) approach to realise new high-speed rail (HSR) lines. This is a paradigm derived from industry and logistics. The adoption of the LARG paradigm in transport systems and railway infrastructures has different implications. The main objective of this paper is to investigate the main implications of the Lean component for realising a new HSR line. The Lean paradigm focuses primarily on cost reduction by eliminating waste and activities that do not generate added value. In the context of High-Speed Rail systems, this approach directly influences both infrastructure design criteria and service planning. It requires comparing the requirements defined during the design phase with the actual, measurable outcomes observed during operation (ex post). This study focuses on the LEAN component applied to a case study relative to the extension of the Italian high-speed rail network along the southern corridor. The case study raises significant questions about the effectiveness of national infrastructure investments, emphasising the need for a broader discussion about the technical criteria used to develop existing high-speed rail lines. This work is of interest to the scientific community, as well as to public and private decision-makers involved in transportation planning connected to high-speed rail infrastructure.

Keywords: Network design; Transport supply · Transport infrastructures · Transport services · Sustainability

1 Introduction

High-speed rail systems are one of the most effective tools for promoting sustainable mobility. Rail transport, in fact, has a lower environmental impact than air travel, especially for medium and long-distance travel. This makes rail transport particularly relevant in the context of the global action of the United Nations, with reference to two Sustainable

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 598–612, 2027.
https://doi.org/10.1007/978-3-032-30530-5_38

Development Goals (SDGs) [1]. SDG 9 "*Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation*" directly concerns infrastructure in its target 9.1, which promotes the realization of "*quality, reliable, sustainable and resilient infrastructure, including regional and transborder infrastructure*"; this aspect adds to the decarbonisation, expressed by SDG 13 "*Take urgent action to combat climate change and its impacts*". The reduced environmental impact in terms of CO_2 emissions, more in general in terms of GHG, and the need to support the economic development of the corridors through infrastructure make rail transport a central element of sustainable development.

The European Union intends to achieve the SDGs of the 2030 Agenda through the development of the Trans European Network – Transport (TEN-T) and the related corridors [2]. In this context, in November 2025, the European Commission presented a comprehensive plan to accelerate the development of the high-speed rail network, with the aim of improving connectivity between major European cities, reducing travel times and strengthening the EU's global competitiveness [3]. The plan defines the reduction of travel times, considering different tools, including the opening of new cross-border connections that will enable rapid connections within individual member states and between different countries. By 2040, the goal is to create a European high-speed network, interoperable and integrated into the EU's multimodal system. The European plan aims to eliminate current infrastructure bottlenecks and harmonise operating rules with a coordinated approach involving all member states.

European perspectives underline the importance of the railway system, and high-speed rail in particular, in achieving economic, social, and environmental sustainability objectives, as well as promoting Europe's territorial cohesion.

This research has the general objective of identifying the main criteria that can guide the planning and design of high-speed railway lines (HSR), solving the programming phase.

The LARG (Lean, Agile, Resilient, Green) approach is one of the possible frameworks for developing new high-speed rail (HSR) lines during the programming [4].

These approaches emphasise the importance of studying transportation demand and related behavioural models that represent user choices in the field of Transportation System Models [5]. Modelling tools are needed to estimate users' travel choices [6] and therefore diverted [7] and induced demand [8]. However, it is necessary to use approaches that focus on the design of the transport supply component [9].

In this context, the main specific objective of this paper is to investigate the main implications of the Lean component for realising a new HSR line. This approach is used in other sectors such as industry and logistics. The Lean paradigm focuses primarily on cost reduction by eliminating waste and activities that do not generate added value. In the context of High-Speed Rail systems, this approach directly influences both infrastructure design criteria and service planning. It requires comparing the requirements defined during the design phase with the actual, measurable outcomes observed during operation (ex post).

The factors that influence the programming of HSR infrastructures include: physical characteristics, with reference to the development of planimetric and altimetric profiles;

technological characteristics, with reference to energy systems; and performance characteristics of high-speed rail services (HST), with reference to routes, travel times and frequencies.

In this context, the Lean component is regarded as a decision-oriented approach in which designing choices are evaluated based on their ability to reduce inefficiencies and enhance system performance. Any infrastructure or network feature that generates unnecessary travel time, indirect routing, capacity mismatches, or discontinuities in service is a waste; improvements in accessibility, travel time reduction, reliability, and effective network integration are values.

Within this perspective, standard railway design and evaluation criteria can be reinterpreted as components of Lean principles. Alignment choices directly affect route directness and operating conditions, thus contributing to the minimization of time-related inefficiencies. Technical Specifications for Interoperability (TSI) ensure system compatibility, continuity, and capacity adequacy, reducing the risk of operational disruptions and underutilization. Network centrality indicators, in turn, measure the ability of the infrastructure to enhance connectivity and accessibility, thereby capturing the generation of user-oriented value.

This step-by-step mapping allows framing conventional engineering criteria within a coherent Lean approach. Rather than representing a separate layer of technical assessment, these criteria constitute the measurable dimension through which waste reduction and value generation are implemented in railway infrastructure planning, consistently with their interpretation in industrial engineering and logistics.

This paper focuses on the Lean component of infrastructure planning, isolating it from other complementary perspectives such as economic appraisal, environmental assessment, or governance-related considerations. This choice is motivated by the need to clarify how Lean principles can be translated into operational criteria for railway network design and evaluation. The paper aims to provide a more explicit and methodologically consistent interpretation of waste reduction and value generation within the specific domain of transport infrastructure and services.

The proposed approach is tested through an application to Southern Italy, a context characterized by heterogeneous infrastructure quality and ongoing high-speed rail developments. This case study is used to evaluate network performance and to identify structural gaps and missing links within the existing railway system.

In accordance with its objectives, the paper, following this introduction, is organized into three sections. Section 2 describes the main materials and methods used to identify the factors of the Lean approach with reference to the physical and performance characteristics of a high-speed railway line. Section 3 presents some examples of the application of the Lean approach to the extension of the HSR network in the case study of the extension of the Italian network towards Southern Italy. Section 4 concludes the paper with a discussion of the results obtained, with some concluding remarks and future research developments.

2 Lean Approach in HSR Programming: Main Factors

According to Campos and De Rus [10], four main models of high-speed rail networks can be identified:

- Model 1 – Exclusive model: High-speed and conventional networks are completely separate, with no interoperability. Services are only possible if both origin and destination are on the same network (e.g., Japanese maglev lines).
- Model 2 – High-speed–mixed model: High-speed trains can run on both high-speed and conventional lines, allowing through-services between the two networks. Conventional trains cannot access high-speed lines (e.g., France and Italy).
- Model 3 – Conventional-mixed model: Certain conventional trains can operate on high-speed lines, enabling through-services, but high-speed trains cannot run on conventional lines. This occurs where different gauges coexist (e.g., Spain).
- Model 4 – Fully mixed model: Full interoperability between high-speed and conventional networks, allowing continuous services across all routes (e.g., Germany).

The implementation of each model requires the specification of the characteristics of the infrastructures (Sect. 2.1) and of the services that travel on them (Sect. 2.2).

2.1 Infrastructure

A high-speed line must be planned, programmed and designed to shorten distances and therefore reduce travel times between the physical locations it connects (two cities or two railway final terminals). International technical and scientific literature identifies the main criteria that guideline planning, programming and design. These criteria concern the geometric, physical, and technological characteristics of high-speed rail lines.

In this context, it is necessary to recall the Technical Specifications for Interoperability (TSI) published by the European Community in 2024 [11]. These are the standard technical and operational characteristics necessary to standardise the railway infrastructures belonging to the Trans European Network – Transport (TEN-T). The TSIs ensure the interoperability of the railway system of the European Union, in relation to the passengers and freight mobility components. The specifications regard different transport system supply components in relation to railway infrastructures, management, and rolling stock. The final aim is to eliminate trade barriers for both industry and consumers within the European Union (EU).

In accordance with the aims of this paper, Table 1 lists the main characteristics of a high-speed railway line indicated with the acronym “INF TSI”. The table shows the TSI codes relating to rail passenger transport with speeds between 120 and 350 km/h. Both “hard” characteristics, like the gauge and the axle load and “soft” characteristics, as the line speed and usable length of platform, are included. The traffic code “P1” indicates the greatest values and then the best performances for interoperability HSR lines. These target values must be considered for realising new HSR lines or renewing existing conventional railway lines.

Furthermore, with reference to the programming and designing of a high-speed railway line, it is necessary to recall the geometric characteristics:

- of the planimetric layout which describes the geometry of the tracks in the plane, composed of straight lines connected by circular curves and transition curves which guarantee gradual variations in curvature and lateral acceleration to increase the safety, comfort and dynamic behaviour of the train; for high-speed lines the planimetric

Table 1. Performance parameters for passenger traffic (source [11])

Traffic code	Gauge	Axle load [t]	Line speed [km/h]	Usable length of platform [m]
P1	GC	17*	250–350	400
P2	GB	20	200–250	200–400
P3	DE3	22.5 (**)	120–200	200–400
P4	GB	22.5 (**)	120–200	200–400

layout guarantees to shorten the distances between the places to be connected in accordance with the horizontal alignment criterion [12] (Fig. 1a);

- of the altimetric layout that defines the vertical direction of the track axis consisting of constant slopes (gradients) connected by vertical connecting curves; the layouts are designed to control variations in gravitational acceleration and limit the longitudinal forces acting on the train; the altimetric layout guarantees optimal conditions of safety, operational efficiency and passenger comfort; along high-speed lines, altimetric variations are significant because they influence the performance of the train and the travel comfort of passengers, in accordance with the vertical alignment criterion [13, 14] (Fig. 1b).

Both criteria aim to ensure that the performance of high-speed rail systems is competitive with alternative transport systems over medium and long distances.

A synthetic measure of horizontal alignment is straightness centrality, which measures how directly a station is connected to the others, and is defined by the ratio between the shortest path and the Euclidean distance [15]. The measurement allows to identify stations with high path linearity, highlighting efficient and direct connections.

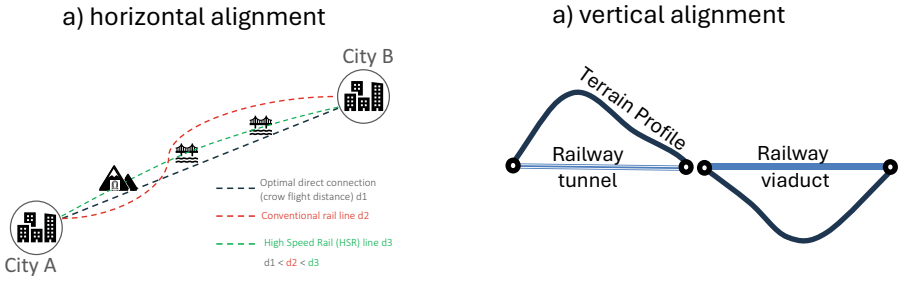


Fig. 1. Horizontal and vertical alignment

Some recent works have analysed ex post the planimetric and altimetric characteristics of the high-speed lines built in the world. Russo et al. [16] analyzed lines built in Italy, France, and Spain, comparing the lengths of existing lines in terms of: the length of conventional lines, the length of high-speed lines, and the straight-line length connecting

the pair of locations (e.g., cities). The authors demonstrate that with advancing technological progress, the length of railway lines tends to decrease asymptotically toward the straight-line length.

However, it should be noted that reducing the length of the tracks from a planimetric point of view could impact on the altimetric characteristics of the route and therefore the vertical alignment. Achieving these standard characteristics impacts the structures to be constructed (e.g., tunnels and viaducts). Following the horizontal and vertical alignment criteria implies the need to construct structures aimed at ensuring straight planimetric routes and low gradients, within the limits indicated by the TSI.

From the point of view of electric power technologies, it should be noted that the hybrid service model implies critical issues in the transition of railway trains from one network to another.

In this paper, the focus is mainly on Model 2 and, to some extent, on Model 4, as classified by Campos and De Rus [10]. The problem of discontinuity in the power supply between networks is therefore central. Although this discontinuity is generally addressed and resolved effectively through electromechanical solutions, this paper aims to highlight the aspects related to the functional components of passenger services, defined in the literature as "hybrid" services.

The interruption of electrical voltage—and therefore of traction power—requires the creation of a short neutral section where trains cannot draw energy. In this phase, the driver (or the on-board system in the case of multi-voltage locomotives) switches from one power system to the other.

Two main problems may arise. The first concerns the train being left without power, even if only for a few moments, resulting in a power outage for on-board equipment with potential repercussions on service efficiency. The second is more general and concerns the reduction in line capacity, due both to switching procedures and to the temporary reduction in power available for traction.

2.2 HST Services

Considering the "full mixed" high-speed rail services [10], the Lean paradigm can be assessed by starting from some fundamental elements characterising the transportation supply. To analyse the rail service supply, a rail service network analysis can be used. The analysis concerns the current structure and an initial assessment of the effects produced by hypothesizing changes. In the discussion below, the four main elements are highlighted sequentially, according to the purposes of the analysis to be conducted.

a) **Travel distances.**

The first element is the distances covered by rail travel. Rail services are used for different trip distances and purposes. Rail services are used for metropolitan and regional connections, with distances of less than 50 km. In this case, rail services compete directly with local or extra-urban road public transport. In the range between 50 km and 600 km, services, generally national, primarily concern connections between very distant centers within the same country or macro-region; competition primarily involves air transport and, in certain geographically defined circumstances, ships. For connections exceeding

600 km, whether national or international, rail services remain competitive compared to air transport, which remains privileged at these scales.

b) **Type of service.**

Once the travel distance scale has been defined, it is possible to identify the main characteristics of the network to be studied. As previously defined, rail services depend largely on the identified distance scale and on competition from other modes of transport. In a general way, high-speed, conventional, and metropolitan services can be identified among railway services, according to operating speeds and frequencies. While metropolitan systems serve travel on a local and regional scale, high-speed and conventional services may vary, considering specific characteristics of the national transport supply system. According to Campos and de Rus [10], four models of high-speed and conventional systems may be defined; this depends on how the trains (HS and conventional) use the infrastructure (HS and conventional). In general, services operated by HS trains are referred to as HS even if they run on a conventional network, as is the case in Italy. In turn, conventional services can be divided based on the scale of the connections (regional, intercity) or the operating conditions (night).

c) **Operator.**

Identifying the operator is necessary to select the transport services operating in the study area. Depending on the distance classes and the type of service, it is possible to identify the railway operators that offer high-speed services. There may be a single railway operator that operates under monopoly conditions on the entire network, or in a liberalized market, there may be multiple railway operators operating on the same origin-destination connections. Operators often coincide with specific administrative extensions and service classes; for example, some operators manage their services on a national railway system; others operate services only within specific classes (e.g., Nuovo Trasporto Viaggiatori covers only HS services in Italy). Once the operators that manage the services identified in the chosen study area have been identified, it is possible to access their respective timetables.

d) **Network nodes.**

From the previous step, it is necessary to filter the relevant nodes within the timetables according to two sequential processes of exclusion and aggregation. These steps depend on the study area, the services chosen, and the type of analysis to be conducted. This choice influences the "granularity" of the study. The exclusion step is motivated by the need to remove stations that are not relevant to the analysis. In a study that focuses on connections between large urban centers on a national scale, some peripheral stations may not be considered. Aggregation is necessary to concentrate attention on a single node, ensuring a functional representation at the scale being analyzed. A national-scale study to evaluate connections between cities implies the need to select the most representative station (for example, the one serving the city centre) to avoid overestimation within the same urban area. The maximum level of disaggregation possible is linked to the individual station if it appears in at least one of the services identified in point b). All possible aggregations, starting from this set of stations, depend on the type of analysis one wishes to conduct and the scale of the study. On a national scale, it is possible to

aggregate and restrict the set to regional capitals; the same aggregation, on a regional basis, would not be possible.

The summary of the relationships between the analysed components is reported in Fig. 2.

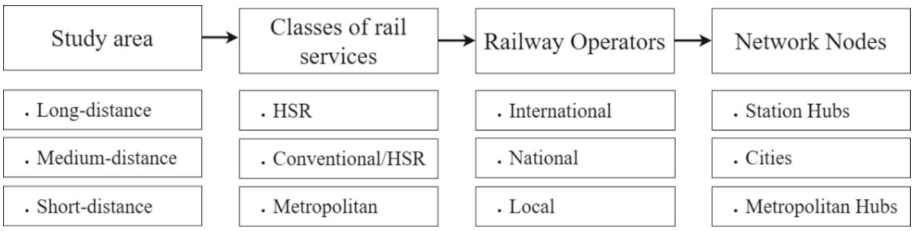


Fig. 2. The four elements to specify for the service analysis

3 Case Studies: HSR Lines for Extension in the South of Italy

The case study concerns the potential development of the high-speed rail network along the TEN-T ScanMed corridor, which connects Italy's southern regions. Currently, the national high-speed network extends to the Campania region, specifically to the Naples Afragola hub. Routes to Bari to the east and Reggio Calabria to the south are currently being programmed and designed. This paper considers the impact on the supply system of the construction of the Naples-Bari and Napoli/Salerno-Reggio Calabria lines.

In the following, in accordance with what was presented in Sect. 2, the infrastructures (Sect. 3.1) and services (Sect. 3.2) are analysed in the current situation and in hypothetical scenarios along the ScanMed corridor.

3.1 Infrastructures

In order to verify the application of the horizontal alignment criteria (Sect. 2.1), the characteristics of some existing high-speed railway lines in Europe were compared. Some lines are present in:

- Italy, considering the six HSR lines that develop from the north (Turin-Milan-Rome) to the south (Rome-Naples/Salerno);
- France, considering the six HSR lines that run from east to west (e.g., Strasbourg-Paris) and from north to south (e.g., Paris-Lyon-Marseille);
- Spain, considering the four lines that run from east to west (Barcelona-Madrid) and from north to south (Valladolid-Madrid-Malaga).

For the purposes of the analysis, the following variables were collected:

- l_{HS}, indicates the linear extension of an HSR line;
- l_{CN}, indicates the Conventional line;
- l_{SL} indicates the Straight line or crow flies line.

The data collected has been analysed and compared. The results of the analyses lead to several considerations. The average length of high-speed rail lines (l_{HS}) in the three countries is about 13% longer than the straight-line distance (l_{SL}) between the terminal stations considered. However, when compared with the average length of conventional rail routes (l_{CN}), the high-speed lines extension shows an average reduction of 15%.

Regarding the verification of vertical alignment, the data on the average gradients of the high-speed railway lines built in the European countries under analysis were compared. It should be noted that the lines built in the three European countries refer to different models of the classification proposed by [10]. The Italian system can carry both freight and passenger trains, and therefore the gradients do not exceed 30%. In France and Spain, the rail infrastructure is dedicated to high-speed passenger trains. These lines are not designed to support heavy freight traffic. Therefore, the maximum gradient is higher, up to 30%, but this allows freight traffic for the shippers.

With regard to the power-supply systems, the existing HSR networks have no homogeneous characteristics between different countries and, in some cases, within the same country. For instance, in the Italian HSR there are nodes powered by 25 kV AC (Reggio Emilia node), where a transition occurs between the 25 kV AC to the 3 kV DC. For instance, at the Napoli Afragola railway station, there is a discontinuity from Roma–Napoli to the Napoli–Salerno lines. In general, there is a high number of nodes directly affected by the introduction of hybrid operations. In general, and regardless of the historical evolution of national electrification systems, lines supplied at 25 kV AC provide superior performance compared with those operating at 3 kV DC.

3.2 HST Services

Regarding the service analysis, the following reports some of the results obtained by applying the study to some variable results. Three indicators derived from the network analysis are primarily analyzed for the resulting network:

- Degree Centrality, measures the number of connections of each node with other nodes in the network. Degree Centrality equal to 1 indicates that the node is connected to all other nodes;
- Strength Centrality, measures the number of connections of each node, weighted by a quantitative value. Considering frequencies as the value, the total daily connections of each node is obtained;
- Eigenvector Centrality, measures the node's score, assigning higher values to connections with nodes with higher scores.

The analysis is proposed at three different territorial scales: national sub-national and regional. This approach is intended to both reflect the hierarchical nature of railway planning decisions and to introduce independent levels of analysis. Each scale addresses a specific decision layer: the national level defines the strategic corridors and long-distance connectivity, the sub-national level refines these structures by considering intermediate territorial systems, and the regional level focuses on network performance in terms of accessibility and internal connectivity in peripheral areas. The three levels are therefore analytically distinct but functionally interdependent, as choices made at the higher scale constrain and inform the evaluation at the lower ones. In this framework, the transition

between scales is both a change in spatial resolution and a shift in the type of performance being evaluated: from corridor efficiency (national level) to territorial coverage (sub-national level), and finally to accessibility and network cohesion (regional level).

National

The first case study involves a national analysis of HS rail services operated by the two Italian HS operators (Trenitalia and NTV). The reference nodes are metropolitan cities, specific territorial entities identified in the Italian administrative system as significant centers of demand generation and attraction. Of the 14 metropolitan cities in Italy, the study analyzed 10 across the peninsula. The study calculated Degree Centrality and Strength Centrality based on the frequencies of the Italian HS network to identify any territorial differences between the ten cities. Compared to the network classification defined in Sect. 2, the Italian system is a Model 2 system, with HS trains able to travel on conventional lines. Therefore, services declared by operators as HS can travel on both HS and conventional lines. Results obtained from the evaluation are reported in Fig. 3. The study shows that the more peripheral centers such as Reggio Calabria and Bari have very low values of Strength Centrality based on frequencies: connections exist but are few. The priority axis of connections is the one that links Milan with Naples, an area served by the HS infrastructure. The Reggio Calabria hub remains largely peripheral, reflecting the lesser development of services in the south, due to the lack of HSR infrastructure. This marginalization of Reggio Calabria's network also determines the peripherality of Sicilian cities; Sicily, about 3 km from the mainland at its narrowest point, is home to a population of over 4.5 million, excluded from the national HSR network [17].

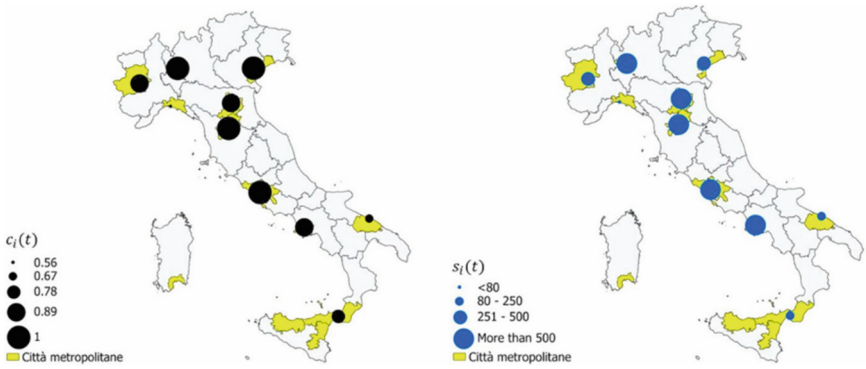


Fig. 3. National. Results for degree centrality and strength centrality. Source: [17]

Sub-national

In this study, the analysis was conducted on a macro-regional scale (less than 500 km), considering the case of central and southern Italy. The services analyzed are HS services, and the two network operators (Trenitalia and NTV) were considered. The considerations to be made regarding the nature of the service are the same as for the National case previously described. It should be noted that, with regards to the study area, only

the connections between Rome and Naples operate on a HS infrastructure, while the remaining part of the analyzed system travels on conventional lines. No aggregation or exclusion was performed; the reference nodes of the network are the railway stations involved in at least one of the services. The main results of the analysis are reported in Fig. 4. The network is divided into two main routes, an Adriatic one centered on Bari and a Tyrrhenian one centered on Naples. Naples and Bari are the two local hubs, while only two stations cover both routes (Fig. 4 right), Rome Termini and Taranto. While Rome has very high values on all indicators, Taranto has low values, linked to the limited presence of connection services. The analysis shows a network centered on Rome, the country's capital, despite its geographic proximity to other centers; even a large city like Naples has very low centrality values [18].

Regional

The third case study concerns a regional-scale analysis, considering the Sicily region in Southern Italy. The island is not currently served by a high-speed rail (HSR) network. However, this characteristic is precisely what makes it relevant within the proposed framework. Sicily represents a context in which HSR is in the planning and implementation phase, limited, for now, at the Palermo–Catania corridor, allowing the analysis to focus not on the performance of an existing system, but on the identification of structural gaps and missing links in the current network configuration. From this perspective, the case study serves two purposes. First, it provides an application of network analysis tools at the regional scale, highlighting deficiencies in accessibility and connectivity that would not emerge from a purely national-level assessment. Second, it allows testing the proposed approach in a "pre-HSR" context, where Lean-oriented principles can support the prioritization of interventions by identifying where infrastructure improvements would generate the highest system-wide value. Sicily is a representative case of regions undergoing transition towards higher-performance rail systems. For this analysis, all conventional services are considered, as there is no high-speed rail network on the island. The operator studied is Trenitalia, which is the only operator on the main standard-gauge regional network. The nodes considered concern the nine former provincial capitals, to capture the relationships between the main centers of generation and attraction of demand. The indicators studied are Strength Centrality and Degree Centrality. The study carried out a comparison with extra-urban bus connections on the same relationships. The main results are reported in Fig. 5. The resulting rail network graph is disconnected, since Trapani, the westernmost city on the network, is excluded from connections with other cities. In general, the bus network dominates the rail network for all stations in terms of degree centrality. Messina is the only city better connected by train than by bus, due to the better rail lines on the network. The result shows a system of connections which, although present, tends to prefer some routes – due to greater demand or better condition of the infrastructure – over others, and to a general lower presence of rail services compared to road services [19].

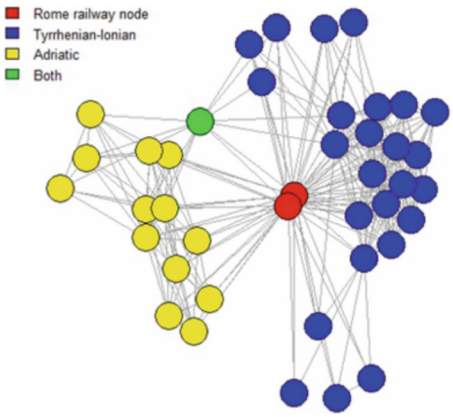


Fig. 4. Subnational. Results for degree centrality, eigenvector centrality and strength centrality (left) and graph representation of relations (right). Source: [18]

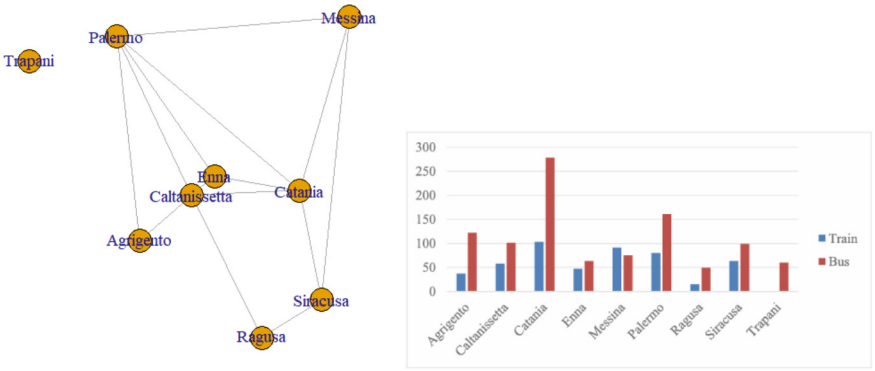


Fig. 5. Regional results. Graph representation (left); Strength Centrality (right). Source: [19]

4 Discussion

Based on the findings of this paper, it appears that the design, programming and planning criteria for HSR infrastructure lines follow specific criteria. The construction of high-speed lines around the world follows a lean approach from an infrastructure perspective through the application of horizontal and vertical alignment criteria.

Consistent with the analysis of lines already built in Italy, France, and Spain, the results show an average reduction in the length of high-speed lines compared to conventional lines of 15%. Continuity between power supplies must also be ensured to limit disruption and ensure high speed performance.

The methodology described for services allows for a comprehensive and concise assessment of rail transport networks, considering four fundamental elements. The

assessments carried out on the network's operating conditions can be combined with scenario hypotheses. Consider the national case of HSR services in Italy. With the current high-speed service network, considering a travel time threshold of 3 h, two southern metropolitan cities, Bari and Reggio Calabria, are disconnected [9]. By adopting a "what if" approach that combines classic transport system models and network analysis, it was possible to compare different intervention scenarios.

Considering the major ongoing projects on the high-speed rail network (Napoli/Salerno-Reggio Calabria, Naples-Bari, and Terzo Valico dei Giovi), the main centrality indicators have improved significantly. This translates into an increase in connections. For example, Reggio Calabria will be connected to other metropolitan cities such as Rome and Bari in a travel time of less than 5 h; with Naples in less than three hours; similarly, Naples and Bari would be connected in less than three hours. These connections depend on the selected program to connect Rome and Reggio Calabria in 3 or 4,5 h. The two programmed options influence the quality of connection rates across the entire Southern Italy network. This would lead to a more even distribution of the opportunities offered by the new HSR connections between metropolitan cities. Furthermore, it would allow Sicily to become closer to the national network; Sicily, in fact, is home to three metropolitan cities, with more than 4.5 million inhabitants.

5 Conclusive Remarks

This paper analysed the Lean component of the LARG paradigm for the construction of high-speed rail lines. The results of ex-post analyses conducted on existing HSR lines were presented. The picture that emerges is not entirely homogeneous, especially across Europe. From an infrastructural perspective, it emerges that the existing lines were built following the principles of horizontal and vertical alignment. Indeed, the extension of the existing lines progressively approaches the distance as the crow flies; furthermore, the vertical gradients do not exceed 30%, also considering the type of traffic that can potentially transit them (passengers/freight).

From a service perspective, it is noted that where pure high-speed systems are present (maximum speed 300 km/h), connection times are significantly reduced, ensuring users can make round trips within the same day.

Although this paper has some limitations, it focuses specifically on one component of the LARG paradigm. The significant role of transport demand and the externalities produced are considered alongside the other components (ARG), also in relation to other medium- to long-distance transport modes. The paper has intentionally concentrated on the Lean dimension of railway infrastructure planning, without addressing other aspects such as economic feasibility, environmental impacts, or institutional constraints. While this represents a limitation, it is also a necessary simplification to make the contribution more clearly identifiable. The results should therefore be interpreted as a methodological exploration of how Lean principles can inform the identification of inefficiencies and the prioritization of interventions, rather than as a comprehensive decision-support framework.

Future research should aim to integrate the Lean-based approach proposed here with more established evaluation methods, in order to assess how efficiency-oriented

criteria interact with economic, environmental, and social objectives in complex planning processes.

The results obtained are consistent with existing studies on railway network performance and accessibility, which highlight persistent disparities between core and peripheral regions. Previous research has emphasized the role of network structure and service levels in shaping accessibility patterns, often identifying similar weaknesses in terms of connectivity and travel times. However, the contribution of this study lies in the interpretation of these findings through a Lean perspective. This shift in perspective allows not only to confirm known criticalities, but also to prioritize them based on their impact on overall system efficiency and value generation.

The analyses and results are nevertheless useful for public and private decision-makers involved in the planning, design, and construction processes of high-speed lines.

Acknowledgments. This study was carried out within the MOST – Sustainable Mobility National Research Center (CUP: C33C22000240001) and received funding from the European Union Next-Generation EU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1033 17/06/2022, CN00000023). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

Credit Author Statement. Conceptualization: C.R and A.R.; Methodology: C.R. and A.R.; Software: A.R.; Validation: C.R. and A.R.; Formal analysis: C.R. and A.R.; Investigation: A.R.; Resources: C.R.; Data curation: A.R.; Writing - Original Draft: C.R. and A.R.; Writing - Review & Editing: A.R.; Visualization: C.R.; Supervision: C.R.; Project administration: C.R.; Funding acquisition: C.R.

References

1. UN: Sustainable Development Goals (SDG). https://sdgs.un.org/goals. Accessed 07 May 2023
2. European Commission: Regulation (EU) 2024/1679 of the European Parliament and of the Council of 13 June 2024 on Union guidelines for the development of the trans-European transport network, amending Regulations (EU) 2021/1153 and (EU) No 913/2010 and repealing Regulation (EU) No 1315/2013 (Text with EEA relevance) (2024)
3. European Commission. Directorate-General for Mobility and Transport: High-speed rail plan. Publications Office of the European Union, LU (2025). https://doi.org/10.2832/0466208
4. Russo, F.: Which high-speed rail? LARG approach between plan and design. Future Transp. **1**, 202–226 (2021). https://doi.org/10.3390/futuretransp1020013
5. Cascetta, E.: Transportation Systems Engineering Theory and Methods. Springer, New York (2013)
6. Russo, F., Musolino, G., Sgro, D.: Estimating run choice models with innovative data collection: day-to-day tickets evolution in high speed rail. Transp. Policy **176**, 103909 (2025). https://doi.org/10.1016/j.tranpol.2025.103909
7. Russo, F., Moschella, M., Musolino, G.: Railway demand evaluation: HSR induced component. In: Gervasi, O., Murgante, B., Garau, C., Taniar, D., Rocha, A.M.A.C., Faginas Lago, M.N. (eds.) ICCSA 2024. LNCS, vol. 14822, pp. 173–187. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-65318-6_12

8. Russo, F., Moschella, M.: Mode choice at national level: revealed and stated preference survey for high-speed rail realization. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15894, pp. 359–374. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97648-3_24
9. Rindone, C., Russo, A.: Centrality analysis for railway network planning: constrained measures for high speed. Transp. Eng. **22**, 100393 (2025). https://doi.org/10.1016/j.treng.2025.100393
10. Campos, J., De Rus, G.: Some stylized facts about high-speed rail: a review of HSR experiences around the world. Transp. Policy **16**, 19–28 (2009). https://doi.org/10.1016/j.tranpol.2009.02.008
11. European Commission: Commission Regulation (EU) No 1299/2014 of 18 November 2014 on the technical specifications for interoperability relating to the 'infrastructure' subsystem of the rail system in the European Union Text with EEA relevance (2014)
12. Li, W., et al.: A method for automatically recreating the horizontal alignment geometry of existing railways. Comput. Aided Civ. Eng. **34**, 71–94 (2019). https://doi.org/10.1111/mice.12392
13. Li, W., et al.: Mountain railway alignment optimization with bidirectional distance transform and genetic algorithm. Comput. Aided Civ. Eng. **32**, 691–709 (2017). https://doi.org/10.1111/mice.12280
14. Li, W., Pu, H., Schonfeld, P., Zhang, H., Zheng, X.: Methodology for optimizing constrained 3-dimensional railway alignments in mountainous terrain. Transp. Res. Part C Emerg. Technol. **68**, 549–565 (2016). https://doi.org/10.1016/j.trc.2016.05.010
15. Das, J., Ram, S.: Exploring the relationship of transport network and land use patterns: an approach through weighted centrality assessment. Urban Plan. Transp. Res. **12**, 2323946 (2024). https://doi.org/10.1080/21650020.2024.2323946
16. Russo, F., Rindone, C., Maiolo, G.A.: High-speed railway planning for sustainable development: the role of length between conventional line and straight length. Future Transp. **5**, 68 (2025). https://doi.org/10.3390/futuretransp5020068
17. Rindone, C., Russo, A.: Intermediate results for railway capacity evaluation: a network analysis for HSR services in Italy. Presented at the AIIT 4th International Conference on Transport Infrastructure and Systems (TIS ROMA 2024), Rome Italy, 19–20 September 2024 (2025)
18. Rindone, C., Russo, A.: A network analysis for HSR services in the South of Italy. In: Gervasi, O., Murgante, B., Garau, C., Taniar, D., Rocha, A.M.A.C., Faginas Lago, M.N. (eds.) ICCSA 2024. LNCS, vol. 14822, pp. 217–232. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-65318-6_15
19. Russo, A., Tesoriere, G., Rindone, C., Campisi, T.: Inter-urban transit connections: a network analysis in an Italian Region. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15894, pp. 375–390. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97648-3_25

Sustainable Tourism Economy, Silk Road, Reuse and Enhancement of Cultural Heritage (STESilkRoad 2026)

Cultural Heritage, Identity, and Power Dynamics: Guiding Practices in Oradea, Romania

Corina-Florina Tătar[1(✉)], Celestina Fazia[2], Clara Stella Vicari Aversa[3], and Klara Szylagyine Czimre[4]

[1] University of Oradea, 410087 Oradea, BH, Romania
[2] Department of Engineering and Architecture, University Kore of Enna, Enna, Italy
Celestina.fazia@unikore.it
[3] Department of Architecture and Design, Mediterranea University of Reggio Calabria, Reggio Calabria, Italy
Clarastella.vicariaversa@unirc.it
[4] Department of Social Geography and Regional Development Planning, University of Debrecen, Debrecen, Hungary
czimrecklara@science.unideb.hu

Abstract. The study highlights how heritage is constructed, contested and interpreted in the multiethnic city of Oradea, Romania. Embedded in theoretical conceptual frameworks related to heritage studies, power dynamics and identity, it analyzes controversies related to the naming and renaming of monuments and public spaces across political regime changes, emphasizing how these practices reflect broader debates about public memory. The conceptual framework is consolidated by empirical data collected from 103 respondents from across 18 countries meant to spotlight the perceptions of political influence and interpretative authority in heritage governance. Oradea is used as a case study due to its multiculturalism and ethnic pluralism of Romanian, Hungarian and Jewish past to show how heritage is preserved, which institutional cultural policies are adopted and how monuments, urban public spaces and street names have been preserved or changed across political regimes changes from communism to democracy. In this scope the role of tour guides practices as cultural brokers who can shape the heritage public understanding is investigated. The article concludes that the mixed approach of political power, cultural pluralism and interpretative attitudes shape Oradea city's multicultural heritage, emphasizing both challenges and opportunities of managing diverse past heritage in contemporary Romanian cities for the capitalization of tourism.

Keywords: cultural heritage · licensing of guides · renaming public spaces · tourist guides' narrative

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 615–629, 2027.
https://doi.org/10.1007/978-3-032-30530-5_39

1 Introduction

Heritage is a strong instrument by which societies build and negotiate their collective identity. It is a representation not only of historical facts but a process of selection and interpretation about which scholars have shown that there are political or cultural-related decisions with reference to what to keep, commemorate or erase as well as legitimacy struggles [27, 33]. In post socialist societies where such political regimes were overthrown there is the need to forge new identity and heritage becomes a contested arena. Romania stands as a relevant case for examining these dynamics. Romania has been under a communist political regime ruled by Nicolae Ceaușescu who targeted to create a labor class standardization and uniformization by demolishing historical neighbourhoods, invoking patriotism to legitimize his authority as well as heritage manipulation [16, 35]. After 1989 Romania has reconfigured its collective memory, eliminating social symbols and renegotiating the importance of public spaces [9]. These cases have been extremely complex in multiethnic societies such as Oradea where Romanian, Hungarian and Jewish history intermingle.

This article investigates how heritage, identity, and power interact in Oradea, focusing on three interconnected dimensions: (1) the theoretical mechanisms through which heritage constructs national identity; (2) the politics of monuments and street renaming in post-communist Romania; and (3) the role of tour guides as cultural brokers who shape public understanding of heritage. By combining theoretical and empirical observations of Oradea, the research widens the spectrum of debates on heritage pluralism, collective memory and urban space governance.

2 Methodology

The study synthetizes key theoretical concepts of Benedict Anderson [1], Eric Hobsbawm [13], Laurajane Smith [27], Michel Rolph Trouillot [33], Pierre Nora [23] which provide the conceptual frame for understanding how heritage can related to identity building and political power. Further references relate to the contributions of Verdery [35, 36], Light and Young [16–18], Czepczyński [7] and Ashworth, Graham and Tunbridge [2] on post socialist collective memory and urban toponymy.

These references contribute to the analysis of communist and post communist heritage practices of Romania. Oradea acts as a case study due to its multicultural history and visible heritage plurality. The analysis examines the Hungarian and Jewish heritage preservation within the city, the street renaming processes after 1989 and tourist guiding dynamics within the local context. The case study approach allows a nuanced exploration of how theoretical concepts manifest in a specific urban environment. This methodology enables a comprehensive examination of heritage as a political and cultural process, integrating theoretical insights with empirical observations.

The study represents a mixed research that combines conceptual analysis with an exploratory empirical component, the former relying on consecrated theory from heritage and collective memory research conceptualizing cultural heritage as a mediated social and political process. For the latter empirical component relying on primary data, which were collected by the means of an online survey designed through Google Forms

software, applied during January–March 2026 to 108 participants from 18 countries, mostly acquainted with the cultural heritage of Oradea, allowed the integration of both local as well as international perspectives. The survey was sent to tourism enrolled students doing their practice of tour guiding at the University of Oradea (Romania); students in architecture from Kore and Mediterranea universities (Italy); as well as other partner countries of the University of Oradea (Romania) providing tourism education as well as staff engaged in the tour guiding activity of Oradea city, Romania. The survey included seven themed items related to the perception of the political influence over managing heritage and guiding practices, completed by two demographic variables about age and origin country. The response rate was of 103 respondents from 18 countries (i.e. Azerbaijan, Benin, Ethiopia, Greece, Hungary, Iran, Ireland, Italy, Kenya, Namibia, Poland, Portugal, Romania, Republic of Kosovo, Somalia, Spain, Togo, Ukraine) (Fig. 1). Most respondents (i.e.35,6%) are aged between 18–29 years, closely followed by of 34,6% respondents aged between 45–59 years; 17,3% are aged over 60 years versus 4,8% aged between 35–44 years. The responses were measured using a Likert scale and descriptively analyzed in line with the theoretical framework. The analysis aims to obtain analytical conclusions rather than statistical generalizations, having acknowledged the convenience sampling limitations.

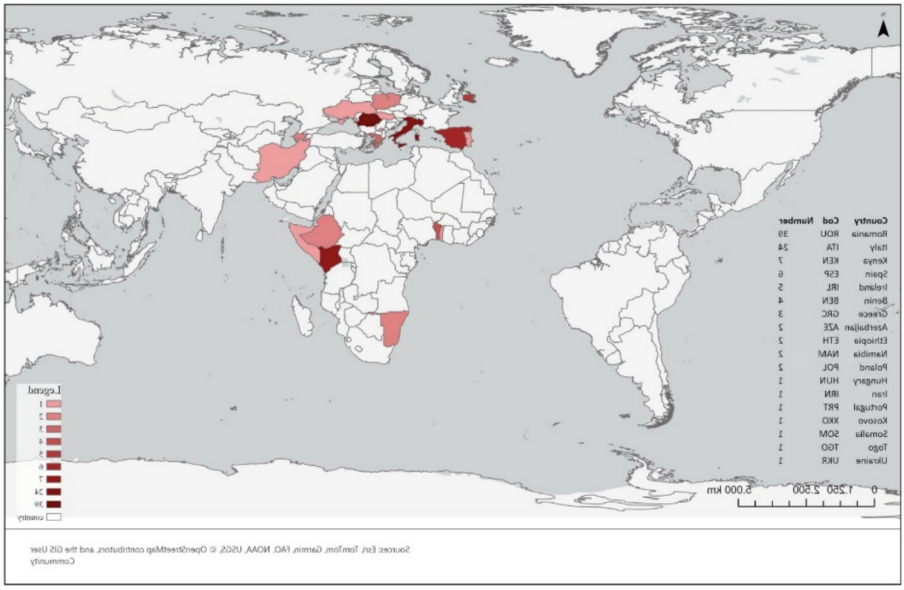


Fig. 1. Respondents' origin country. Source: Authors, based on the survey applied in January–March 2026

International respondents were included because heritage interpretation—especially in tourism and guided tours—is shaped not only by local memory politics but also by the expectations, perceptions, and interpretive frameworks of global audiences. Since

Oradea's heritage is consumed, evaluated, and narrated to transnational tourists, gathering perspectives from respondents outside Romania strengthens the study's ability to analyze how heritage narratives travel, are understood, and are contested across cultural contexts.

3 Theoretical Framing – Heritage, Power and Identity

3.1 Heritage as a Tool for National Identity

Heritage plays a central role in the construction of national identity. Benedict Anderson [1] famously described nations as "imagined communities," held together not by face-to-face interaction but by shared symbols, narratives, and rituals. Heritage—monuments, museums, commemorations, and preserved sites—provides the material and symbolic infrastructure through which these imagined communities are sustained. Hobsbawm and Ranger [13] argue that many traditions presented as ancient are in fact modern inventions designed to legitimize political authority. States selectively elevate certain historical figures, events, and cultural practices to create a sense of continuity and unity. This process often involves the suppression or marginalization of alternative histories that do not fit the desired national narrative.

Trouillot [33] underlines the fact that historical production is shaped by power at each step, from the creation of archives to the construction of collective memory. Laurajane Smith's [27] concept of "Authorized Heritage Discourse" continues to demonstrate how official heritage narratives privilege the elites' monumental and majority cultural expressions while it marginalizes the minorities' voices.

Ceaușescu's Romania provides an example of heritage manipulation. The regime has demolished historical neighborhoods, including churches and synagogues meanwhile building a monumental socialist architecture such as the House of the People [16, 35], advocating for the socialist state. This conservation and selective erasure exemplify how heritage can be used as an instrument to build a politically useful past.

3.2 Controversies Around Monuments and Renaming Public Spaces

Monuments and place names shape collective memory, thus according to Pierre Nora [23] heritage places can stabilize certain narratives of the past and illustrate the era that produced them.

Young [39] debates on a renegotiation of collective memory (rather than its erasure) when heritage such as monuments or street names of oppressive and authoritarian leaders are removed. Trouillot's analysis [33] over history's silence underlines the fact that public commemorations invoke both remembrance as oblivion.

In post-communist Romania, the renaming of streets and removal of socialist monuments have reflected the will to cleanse the public space of discredited symbols [9]. Light and Young [17] show that the name of streets during socialism served as ideological tools integrating the regimes' values into day-by-day life and that 1989 regime change to democracy, cities aimed to restore pre-communist names or introduce new ones, aligned with the national and democratic narratives. As Oradea is a multiethnic city, the renaming initiatives had to navigate through Romanian, Hungarian and Jewish histories, therefore intermingling cultures that are currently capitalized for tourism.

3.3 Power Dynamics Between Majority and Minority Cultures

Heritage may often replicate power hierarchies, and an Authorized Heritage Discourse Smith's [27] may privilege majority narrative as the genuine embodiment of the national identity and minority cultures can be shaded or marginalized. Linda Tuhiwai Smith [28] shows how indigenous cultural practices are often reformulated through colonial or nationalist lens, devoid of political meaning and transformed into consumable symbols. Ashworth, Graham and Tunbridge [2] claim that multicultural societies must admit the past's plurality embodied into their landscape. Not observing this aspect can lead to a "dissonant heritage" in which conflicting nations coexist in tension.

Even UNESCO-labelled heritage designations, often seen as empowering, can consolidate asymmetries by freezing traditions into static forms or privileging majority interpretations [4]. These dynamics highlight the fact that heritage is not only a record of the past but rather a power arena where cultural domination is asserted and contested.

4 Heritage Politics in the Multicultural City of Oradea

4.1 Managing Majority and Minority Heritage

Oradea is a city in which Romanian, Hungarian and Jewish history meet. Unlike many post-socialist cities which attempted to homogenize their heritage, Oradea adopted a heritage pluralism model. This approach is in line with Ashworth, Graham and Tunbridge's [2] argument according to which multicultural cities should treat heritage as a "shared resource".

Oradea city has kept Hungarian monuments and statues, preserved the name of Hungarian historic streets, has restored Jewish synagogues such as the Zion Neologue Synagogue from Independentei street and the Orthodox Synagogue from the Mihai Viteazu street. These efforts ensure that minorities' history remains visible within the urban landscape. Furthermore, Oradea provides a complete education pathway in Hungarian language, from kindergarten to university studies showing an engagement and support towards cultural living practices, not only of symbolic heritage.

Oradea benefits of a wide multiethnic and pluri-confessional cultural heritage of Romanian, Hungarian and Jewish origin among which the later translate in the territory into synagogues and residential and commercial buildings referred to as palaces given their monumental size. Most Art Nouveau buildings in Oradea were the property of the rich bourgeoisie of Jewish origin [12]. But unfortunately, out of 15040 Jews that could be counted at the flourishing economic and artistic boom of Oradea in 1910 [31], in 2025 there are 134 persons of Mozaic confession [11].

The city of Oradea boosts a prolific cultural production from the fin-de-siècle and debut-de siècle Art Nouveau period with 253 recorded historical monument buildings according to the List of the Historical Monument Buildings of Oradea in 2004 [12]. Besides architectural production, the city cultural output was highly intensive with more than 24 existing printing houses at around the year 1900, a trendy lifestyle that attracted tourists who could be accommodated and fed within more than 123 touristic units. Following this development trend, in 1903 the city was electrified with public lightning

and the first functional tram lines in 1906 so that urbanistically the city gains its status of a reputed western city during 1985–1915 [3].

Given this multifaceted heritage landscape with the city's tangible and institutional commitments to heritage pluralism, it is unsurprising that respondents express nuanced views regarding the political dimensions of heritage governance, as reflected in the survey results. Therefore the city's commitment to preserving Romanian, Hungarian, and Jewish heritage side by side fosters a public awareness that recognizes political influence without reducing heritage to a battleground of absolutes.

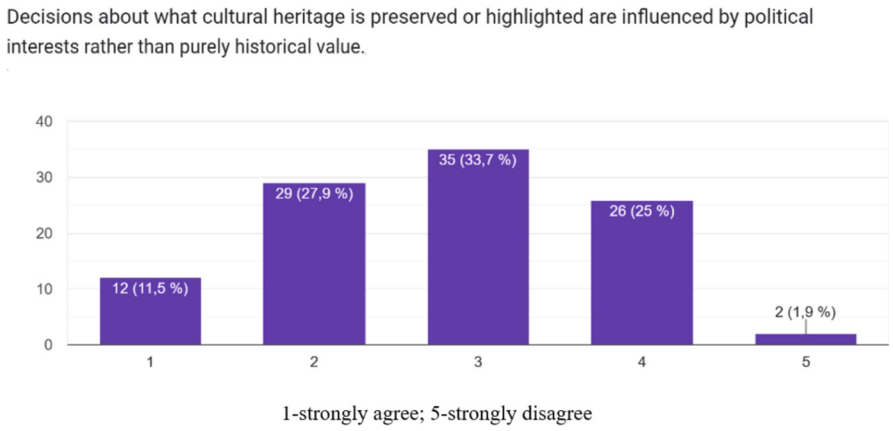


Fig. 2. Respondents' survey results. Source: Authors, survey applied during January–March 2026

The distribution of responses shows that most participants acknowledge some political influence on cultural-heritage decisions, with answers clustering around the mid-range rather than the extremes (Fig. 2), indicating that respondents consider heritage governance as a nuanced process, both shaped by historical as well as political factors. The dominance of the 2–4 category shows ambivalence rather than polarization lining with the academic debates according to which heritage is inherently political [10, 27]. The small number of agreements or strong disagreements continues to advocate for the idea of contested perceptions, but not absolute according to Tunbridge and Ashworth's concept of "dissonant heritage" [34].

4.2 Renaming Public Spaces in Oradea

The remaining of streets after 1989 illustrates how collective heritage is renegotiated after the political regime's change to democracy. The erasing of socialist names and renaming heritage monuments and streets after pre-communist figurehead names took place concomitant with debates concerning the Hungarian cultural landmarks. Czepczyński [7] describes it as the "politics of cultural landscape" in which concurrent groups try to inscribe their identities in the urban space.

Verdery [35] notes that post socialist memory politics often involve struggles over legitimacy, with different communities mobilizing history to assert belonging. In Oradea,

the coexistence of Romanian and Hungarian street names reflects an ongoing negotiation of identity in a multicultural context.

Thus the 23 August Park, as it was named during the communist regime of 1950–1989, which stood for the Communist Party celebration day through chants for the leader was renamed after the Revolution of 1989 when Romania turned to democracy into 1st December Park (i.e. date related to Transylvania, Banat, Crisana and Maramures regions' Union Day with the Romanian Principalities which occurred on 1st December 1918). The currently-named Balcescu Park was named in the 1900s the Rhédey Park [8] when it was under Hungarian occupation. The currently-named I.C. Bratianu Park was named during the communist years 1950–1989 Laborers' Park. On the other hand, Hungarian-named parks can be found in Oradea as the Petöfi Sándor named after the Hungarian revolutionary poet Petöfi Sándor, indicating Oradea's city openness toward multiethnic cultural values.

Many public squares of Oradea bore Hungarian, German or Jewish names in the 19th and early 20th centuries, also known as the Belle Epoque time, such as the pedestrian and currently named street Vasile Alecsandri which was called Grune Baum Gasse; the Herrenstrasse which is currently the Archbishop Roman Ciorograriu Street; the currently-named Union Square which was named Kispiac ter during the years 1714–1740; the currently-named pedestrian Republicii street bore multiple names as Nagy Utca (Great Street) in 1927, Fo utca (Main Street) in 1890; Racoczy in 1900 and Ferdinand Street in 1920 [24]. Not only squares bore the pledge of the historic times they witnessed but also individual attractions or hotels such as Pannonia Hotel in 1904, which was renamed simply Palace during the interwar period and since 1944 Transilvania Hotel. Many attractions from the historical centre of Oradea, which survived the large-scale demolition campaigns of the communist period (1950–1989) aimed at creating standardized workers' housing blocks designed to increase residential density and impose urban uniformity and standardization have been accredited their initial names in the present time, be they Hungarian, German or Jewish-stemmed, so from this point of view we can refer to a heritage positive discrimination from the city's governance. Oradea is a border multiethnic and multicultural city and this influence is strongly visible in the historical city of Oradea, it's as if the historical periods are printed on the buildings' facades of Baroque, Art Nouveau, Neo-Brancovenesc/Romanian and eclectic styles.

Most of the cultural heritage of Oradea bear Hungarian or Jewish names due to the past Austro-Hungarian occupation, which were not changed after Romania passed from a socialist regime to democracy on 22nd December 1989, therefore some examples of palaces or architectural buildings bearing Hungarian or Jewish names, taken after their initial property owners refer to Moskovits Miksa Palace, Rimanoczy Kalman senior Palace, Vigado Halls of the Black Eagle Palace, Ulmann Palace, Bartsch Palace, Szechenyi Hotel, Adalbert Racs House [25].

Furthermore Romania passed a property restitution law in 2000 (Parliament of Romania, 2000) through which all former landlords or their heirs regained their properties taken by the socialist state during 1947–1989, among which in the case of Oradea city, there were many important heritage buildings such as the most monumental heritage building for Oradea: the Baroque Palace which hosted the Crișana Land Museum during 1971–2006 was transferred to its former owners and is currently the headquarters of

the Roman-Catholic Bishopric. The Baroque Palace was built during 1761–1777 by the Viennese architect Franz Anton Hillebrandt, is U-shaped with a façade of 102,3 m^2, hosting over 100 rooms [24].

The built heritage charm and buildings' monumentality of Oradea city is given by the fact that it witnessed an architectural boom during the Belle Epoque area which overlapped the beginning of the 20th century alongside a social, economical and cultural development. Trade owners would display their wealth and status through their residences, some even putting their monograms on the building facades (i.e. Fuchsl Palace) or signs of the trade they professed (i.e. Moskovits-Miksa Palace) as a sign of distinction. Therefore the real estate is booming with administrative, cultural and trade edifices, witnessing an increment in the rented palaces and villas, which were designed by local or Austro-Hungarian architects [25].

The public organization named the Foundation for the historical Monuments Protection was set up in 2011 and administers and uses the brand Oradea Heritage to promote its heritage. It coordinates restauration projects, administers historical buildings, transforms buildings into dynamic cultural spaces, organizes guided tours, exhibitions, cultural events and promotes Oradea heritage nationally and internationally.

5 Guiding and Heritage Interpretation

5.1 Tour Guides as Cultural Brokers

Tour guides can play a crucial role in shaping tourists' vision about heritage. Cohen [6] describes guides as "cultural brokers" who mediate between tourists and local communities and whose narratives can strengthen or challenge the official discourses about heritage.

In Oradea, where Romanian, Hungarian and Jewish histories coexist, guides can underline different aspects of the past according to their own cultural context. This selective framing can influence the way in which visitors perceive identity and local heritage. Trouillot's [33] argument according to which historical narratives are vulnerable to manipulation is very relevant here, so that without proper regulation, guides detain a significant interpretative authority shaping public memory in ways which can support or suppress heritage pluralism. This can particularly resonate with the case of Oradea as it is located at a border area, only 10 km away from Hungary, prone to the distance-decay effect [38], so that tourists from the neighboring Hungarian country have been recorded constantly throughout the city and its contested past may influence foreign unlicensed tour guides' potentially biased narrative. Romania does not have a tourist police to regulate the tour guiding activity and this oversight can allow these transgressions. These practices can consolidate the majority's narratives or fuel a "dissonant heritage".

The chart illustrates that most respondents tend to agree that foreign unlicensed tourist guides can increase the risk of a favoring or unproper heritage interpretation, 34,6% choosing "1" and other 23,1% choosing "2" (Fig. 3). This general trend reflects the public concern about heritage mediators' feasibility and responsibility. Such perceptions align with those of the researchers who claim that heritage interpretation is shaped by authority, education and power dynamics [10, 27]. A relatively small share of

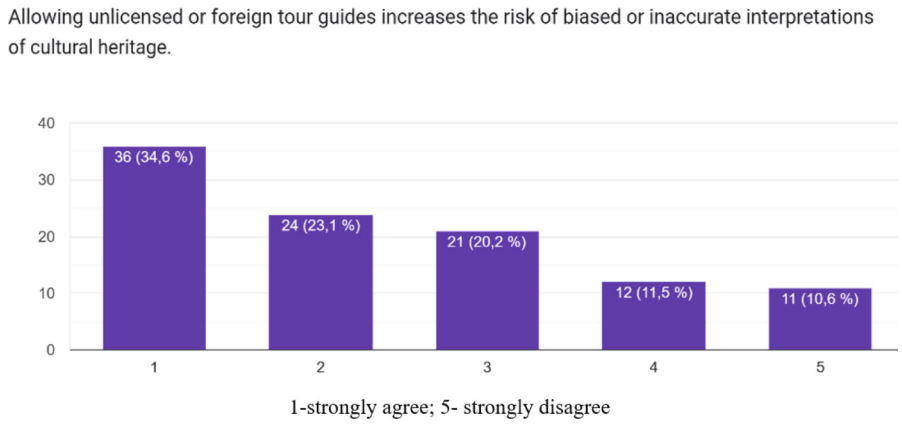


Fig. 3. Respondents' survey results. Source: Authors, survey applied during January–March 2026

strong agreement (10,6%) suggests that despite the admitted risks, they are not considered absolute, leading to wider debates on expertise and authenticity related to heritage tourism [32].

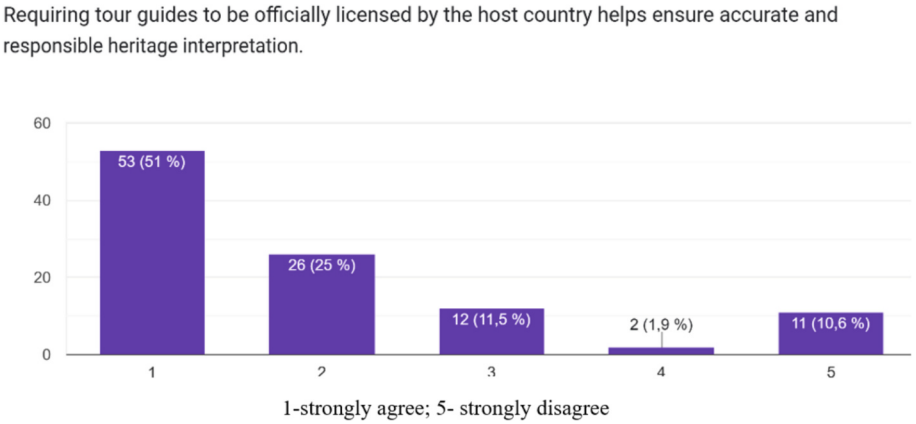


Fig. 4. Respondents' survey results. Source: Authors, survey applied during January–March 2026

The overwhelming support for mandatory licensing reinforces established research showing that trained, officially accredited guides play a crucial role in ensuring accurate, trustworthy (Fig. 4), and professionally accountable heritage interpretation [27, 32].

The concentration of responses around the middle and higher categories reflects widespread awareness that state licensing can shape heritage narratives (Fig. 5), echoing scholarly critiques of how official regulation often reinforces dominant political interpretations [10, 27, 34].

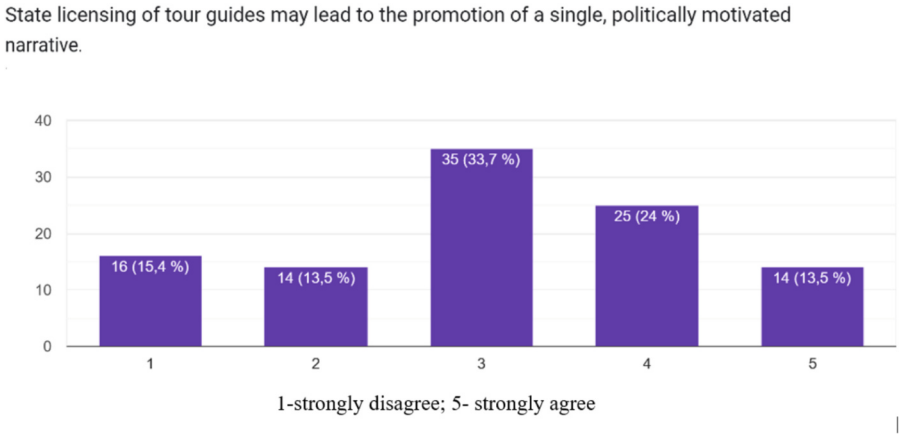


Fig. 5. Respondents' survey results. Source: Authors, survey applied during January–March 2026

5.2 Legal Framework

In Romania the tourist guides' evolution in time is incremental, from merely 24 tourist guides in 1998 to 4,335 licensed tourist guides in 2016 [29] with a high with a high prone towards sustainability practices of facilitating contacts between tourists and locals [30].

At the level of Romania the licensing of tour guides is done according to the Law order 180/2023, modified into 3397/2024 and published in the Official Gazette [15] regarding the conditions and criteria for the selection, schooling, certification and use of tourist guides by the Ministry of Economy, Entrepreneurship and Tourism. All the methodological norms are stipulated with this law order alongside the final dispositions at Chapter V, Article 19 which require that all tourist guides should wear the badge and make proof of the certificate that can be checked anytime by the competent authorities. Within the same article, at number (2) it is mentioned that all foreign tourist guides who are temporarily or occasionally guiding in Romania are not allowed to interpret within the tourist attractions and should appeal to attested local guides. The sanctions are not stipulated within the law and we deem it leaves room for foreign guides to interpret the cultural heritage from the outdoors as well as history based on their good or least good intentions, especially in the case of contested borderland spaces which were under subsequent occupations across time. In order to become a tourist guide in Romania, there's a qualification course that should be attended by candidates with a minimum Baccalaureate diploma, after which a certificate is obtained and eventually the tourist guide badge which should be renewed every 5 years. The latter entitles graduates to legally function as tourist guides, professions categorized into local, national or specialized tourist guide.

The survey's major trend towards an agreement that non-local or international guides can enrich heritage interpretation supports the idea that the strict licensing framework of Romania – despite regulating access to indoor attractions - should continue to admit the added value of diverse external perspectives, mainly in areas with a contested heritage (Fig. 6).

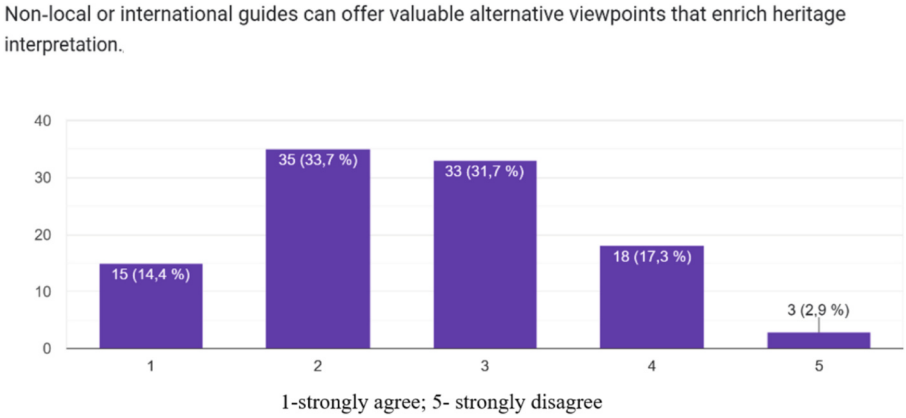


Fig. 6. Respondents' survey results. Source: Authors, survey applied during January–March 2026

5.3 Tour Guiding Practices in Oradea City

At the level of Oradea city there is a legal entity that manages and promotes tourism, known as APTOR (i.e. The Association for the Promotion of Tourism in Oradea and its Region) and it operates the "Visit Oradea" brand, the city's official tourist information platform. Within their activity report for 2025 [37] it is stated that for the year 2025 they provided gratuities and benefits for the tourists visiting Oradea through their 5th campaign if the tourists are accommodated within a rated tourist accommodation unit for a minimum of two consecutive nights such as: gratuities to all museal sections pertaining to the Cris Land Museum – Museum Complex and Bihor Heritage Protection Foundation; 24-hour free transport within all the transport network; discounts to partner cafes and restaurants. This campaign yielded very good results such as a minimum stay of 4,05 nights and 86,09% domestic tourist as well as 13,91% foreign tourists whose origin country is mainly from Hungary, followed by Germany, Poland, USA, France, Slovakia, the Netherlands, Israel, Moldavian Republic and Austria. Thus as a tourist destination, in 2025 Oradea consolidated its 8th rank at the level of the country, as in the previous two years of 2023 and 2024 with 220,277 tourists. Though APTOR, Oradea promotes itself as a surprisingly accessible and unique destination for spontaneous city breaks. It is a European city with Art Nouveau architecture, active cultural life and a slow living urban environment. It advertises itself in the social media through Facebook, Instagram, Tiktok, DOOH campaigns on digital Metrorex screens in Bucharest and collaborations with influencers. APTOR's video postings advertising campaigns triggered many online views of 4.766.838 on Instagram and 5.037.800 views on Facebook (Visit Oradea, 2025).

Part of the same advertising strategy, Visit Oradea has initiated in 2022 the project of providing free guided tours which has uninterruptedly continued until 2026 and which recorded in 2025 a number of 1854 attendees eager to discover the history and architecture of Oradea. In the same year the night guided tours were added, amounting to 54 guided tours that occurred during 30th May–14th September 2025. These tours were provided by licensed tour guides, every Friday, Saturday and Sunday, and upon request also in English and Hungarian languages. Furthermore, in 2025, Visit Oradea in partnership with Spark Communication continued to promote the city for visitation through the

Oradea City Hunt experience to which 323 participants were recorded within 17 tours. Oradea City Hunt is a project meant to target people who want to discover Oradea's hidden gems in an interactive manner.

The strong public agreement that tourists benefit from hearing multiple perspectives, as shown in the graph (Fig. 7), directly reinforces Visit Oradea's strategy of offering diverse, licensed, and multilingual guided experiences—ranging from free historical tours to night tours and the interactive City Hunt—which collectively broaden the interpretive voices available to visitors.

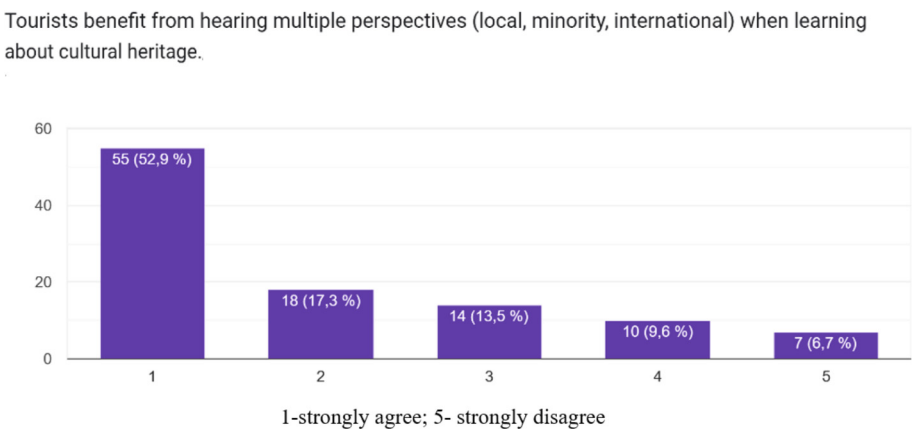


Fig. 7. Respondents' survey results. Source: Authors, survey applied during January–March 2026

Limitations concerning the study relate to the sample's representativeness given that most respondents are from Romania and Italy which might limit results' generalization to diverse tourist populations who are lower represented in the study, as well as limitations related to Oradea's local context. Notwithstanding, the research was meant to gain analytical insights rather than statistical generalizations.

6 Discussions and Conclusions

Heritage in Romania, and particularly in Oradea is shaped by complex interactions between political power, cultural pluralism, and interpretive authority. Theoretical frameworks from Anderson [1], Hobsbawm [13], Smith [27], Trouillot [33], and others reveal how heritage constructs national identity through selective preservation and erasure. Post communist memory politics further complicate these dynamics, as societies renegotiate public memory after regime change.

The results of the survey confirm theories according to which cultural heritage is perceived as a negotiated political process and less an instrument of political absolute domination supporting theories connected to authorized heritage discourse [27] and dissonant heritage [34] applied in an urban multicultural southeastern European context. The study bridges the gap between heritage studies, memory politics and tour guides' practices, an often marginalized actor in the heritage literature. Whereas most studies

focus on monuments, the state and urban policies, the current study shows that guiding interpretation is key for exercising symbolic power. The respondents of the survey admit the existence of a political influence over heritage, and express a rather nuanced positioning. As far as tour guides' activity is concerned the respondents show high credibility in licensed guides and the risk perception associated with unlicensed guiding as well as a conditioned acceptance of foreign perspectives. The results support studies of tour guides as mediators between heritage and consumers [6], unlicensed tourist guiding that can create subjective narratives [33] and field expertise legitimating public trust into heritage interpretation. Tourist police which Romania does not currently have could exercise control and fill the gap of the discontinuity between regulation and practice at an institutional level. Compared to other European Union member countries, a union which Romania joined in 2007, the existence of a tourist police or proposed temporary/specialized police bodies exist in countries such as Spain, Italy, France, Portugal, Greece, Croatia and Bulgaria especially in areas with an increased tourist flow and is meant to assist tourists [22], prevent crime rate [5, 20], prevent abuse and control guides' activity [14, 20].

The approach of Oradea city to heritage highlights both the challenges and opportunities of managing a multicultural history. City heritage preservation of Hungarian and Jewish values, besides the Romanian ones, reflect an engagement towards heritage pluralism. Notwithstanding, the lack of tourist police increases vulnerabilities [22] as guides can deliver selective or biased narratives which challenge official efforts of promoting inclusion.

Overall, Oradea shows how tourist heritage can function simultaneously as a resource for inclusion and a site of contestation. Understanding these dynamics is essential for developing policies that promote cultural recognition, social cohesion, and historically grounded dialogue in multicultural societies. The case of Oradea city offers a compelling model for how Eastern European cities can navigate layered identities while fostering responsible and pluralistic heritage interpretation.

Acknowledgments. We thank all partner universities for sharing the survey link to tourism enrolled students and partner private institutions or staff involved in the tour guiding activity.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Anderson, B.: Imagined Communities: Reflections on the Origin and Spread of Nationalism. Verso, London (1983)
2. Ashworth, G.J., Graham, B., Tunbridge, J.E.: Pluralising Pasts: Heritage, Identity and Place in Multicultural Societies. Pluto Press, London (2007)
3. Borcea, L.: Memoria Caselor/The Memory of Houses, Oradea, pp. 12–13 (2003)
4. Bortolotto, C.: The French inventory of intangible cultural heritage: domesticating a global paradigm into French Heritage Regime. In: Bendix, R., et al. (eds.) Heritage Regimes and the State, pp. 265–282. Universitätsverlag Göttingen (2010)
5. Buil-Gil, D., Mawby, R.I.: Do tourists report crime to the police? An exploratory analysis in Barcelona. Curr. Issue Tour. **26**(18), 2931–2947 (2022)

6. Cohen, E.: The tourist guide: the origins, structure and dynamics of a role. Ann. Tour. Res. **12**(1), 5–29 (1985)
7. Czepczyński, M.: Cultural Landscapes of Post-Socialist Cities: Representation of Powers and Needs. Ashgate (2008)
8. Filimon, C., Tatar, C.F., Filimon, L.A., Herman, L.M.: Urban green spaces – a support for physical activities and tourism. Case study Oradea Municipality, Bihor, Romania. Balt. J. Health Phys. Act. **13**(Spec.Iss.1), 27–39 (2021)
9. Foote, K.E., Tóth, A., Aradi, P.: Hungary after 1989: inscribing a new past on place. Geogr. Rev. **90**(3), 301–334 (2000)
10. Harrison, R.: Heritage: Critical Approaches, 1st edn. Routledge (2013)
11. HCL, HCL-15-25 Hotarare privind aprobarea Planului de Ordine şi Siguranţă Publică al municipiului Oradea pentru anul 2025/ Decision regarding the approval of the Public Order and Safety Plan of the Municipality of Oradea for the year 2025. https://mol.oradea.ro/storage/47706/HCL-415_25.pdf. Accessed 04 Jan 2026
12. Herman, G.V., Tatar, C.F., Puscas, B.D.: Assessing the importance of the Jewish urban cultural heritage in Oradea, Bihor County, Romania, Forum geografic. Studii și cercetări de geografie și protecția mediului **XVII**(**2**), 151–158 (2018)
13. Hobsbawm, E., Ranger, T.: The Invention of Tradition. Cambridge University Press (1983)
14. Kaszás, N., Keller, K.: The emergence of safety and security in the tourism strategies of EU member states. GeoJ. Tour. Geosites **45**(4spl), 1717–1725 (2022)
15. Legislative Portal. https://legislatie.just.ro/Public/DetaliiDocumentAfis/291733. Accessed 10 Jan 2026
16. Light, D., Young, C.: Reconfiguring socialist urban space: the 'systematization' of Bucharest. Hum. Geogr. J. Stud. Res. Hum. Geogr. **4**(1), 5–16 (2010)
17. Light, D., Young, C.: Toponymy as commodity: exploring the economic dimensions of urban place names. Int. J. Urban Reg. Res. **39**(3), 435–450 (2014)
18. Light, D., Young, C.: Local and counter-memories of socialism in post-socialist Romania. In: Beyen, M., Deseure, B. (eds.) Local Memories in a Nationalizing and Globalizing World. Palgrave Macmillan, London, pp. 221–243 (2015)
19. Mawby, R.I., Tecău, A.S., Constantin, C.P., Chițu, I.B., Tescaşiu, B.: Addressing the security concerns of locals and visitors for the sustainable development of tourist destinations. Sustainability **8**, 524 (2016)
20. Mawby, R.I., Vakhitova, Z.I.: Researching the relationship between tourism, crime and security: the tourism industry and the disenfranchised citizens. In: Gill, M. (ed.) The Handbook of Security, pp. 581–601. Palgrave Macmillan, Cham (2022). https://doi.org/10.1007/978-3-030-91735-7_27
21. Murat Payam, M.: Tourism oriented policing and protection services (TOPPs). In: Eurasia International Tourism Congress: Current Issues, Trends, and Indicators (EITOC-2015), pp. 1–7 (2015)
22. Murat Payam, M.: Tourists' security: the need for tourism police in Bosnia and Herzegovina. J. Econ. Manag. Trade **12**(1), 1–9 (2015)
23. Nora, P.: Between memory and history: les lieux de mémoire. Representations **26**, 7–24 (1989)
24. Oradea Heritage: Povestea palatelor oradene- Volumul I/The Story of Oradea's Palaces-Volume 1, Treira, Oradea (2024)
25. Oradea Heritage: Povestea palatelor oradene- Volumul III/The Story of Oradea's Palaces-Volume 3, Treira, Oradea (2020)
26. Parliament of Romania, 2000, Legea 1/2000 pentru reconstituirea dreptului de proprietate asupra terenurilor agricole şi celor forestiere, solicitate potrivit prevederilor Legii fondului funciar nr. 18/1991 şi ale Legii nr. 169/1997 / Law No. 1/2000 on the reconstitution of the right of ownership over agricultural and forest land, claimed in accordance with the

provisions of the Land Fund Law No. 18/1991 and Law No. 169/1997. https://edirect.e-guvernare.ro/Uploads/Legi/15226/LEGE%20nr.%201%20din%202000%20pentru%20reconstituirea%20dreptului%20de%20proprietate%20asupra%20terenurilor%20agricole.pdf. Accessed 15 Dec 2025
27. Smith, L.: Uses of Heritage, p. 368. Routledge, London (2006)
28. Smith, L.T.: Decolonizing Methodologies: Research and Indigenous Peoples. Zed Books, University of Otago Press (1999)
29. Tătar, C., Herman, G., Giurgiu, A.: Analysis of the licensed tourist guides' evolution of Romania and critical issues affecting their activity. Ann. Univ. Oradea Ser. Geogr. (2) (2017). year XXVII
30. Tatar, C., Herman, G., Gozner, M.: Tourist guides' contribution to sustainability in Romania. GeoJ. Tour. Geosites **2**(**1**), 282–287 (2018). Year XI
31. Tikvah, Jews in Oradea (2026). https://www.tikvah.ro/ro/evreii-in-oradea/istoria. Accessed 02 Feb 2026
32. Timothy, D.J., Boyd, S.W.: Heritage Tourism. Prentice Hall, Harlow (2003)
33. Trouillot, M.R.: Silencing the Past: Power and the Production of History. Beacon Press (1995)
34. Tunbridge, J.E., Ashworth, G.J.: Dissonant Heritage: The Management of the Past as a Resource in Conflict. Wiley (1996)
35. Verdery, K.: National Ideology Under Socialism Identity and Cultural Politics in Ceaușescu's Romania, University of California Press Berkeley, Los Angeles, Oxford (2019)
36. Verdery, K.: The Political Lives of Dead Bodies: Reburial and Postsocialist Change. Columbia University Press (1999)
37. Visit Oradea, Prezentarea raportului de activitate/Activity Report Presentation. https://oradea.ro/wp-content/uploads/2026/02/raport-de-activitate-2025-APTOR-prezentare-17.02.26.pdf. Accessed 12 Feb 2026, 25 Oct 2025
38. Yan, L.: Distance decay. In: Jafari, J., Xiao, H. (eds.) Encyclopedia of Tourism, p. 292. Springer, Cham (2024). https://doi.org/10.1007/978-3-030-74923-1_358
39. Young, J.E.: The Texture of Memory: Holocaust Memorials and Meaning. Yale University Press (1993)

Efficient Use of Natural Tourism Resources Within Sustainable Tourism Along the Silk Road: Forecasting Tourist Flows in Uzbekistan Based on the ARIMA Model

Bahodirhon Safarov[1,2], Mansurova Nargiza[2(✉)], and Mustafokhon Toshkhujaev[3]

[1] Samarkand Branch of Tashkent State University of Economics, Samarkand 140147, Uzbekistan

[2] "Silk Road" International University of Tourism and Cultural Heritage, Samarkand 140104, Uzbekistan

n.sh.mansurova@gmail.com

[3] Samarkand State University named after Sharof Rashidov, Samarkand 140104, Uzbekistan

Abstract. This study evaluates the efficiency of natural tourism resource use in Uzbekistan from the perspective of sustainable tourism development. The central premise of the research is that the dynamics of tourist flows reflect not only the overall scale of tourism activity, but also, indirectly, the level of utilization and economic significance of natural tourism resources. For this purpose, the total number of tourists was analyzed using annual time-series data for the period 2010–2024. This indicator was constructed as the sum of domestic tourists and inbound foreign visitors. To identify the future dynamics of tourist flows, the AutoRegressive Integrated Moving Average (ARIMA) model was employed. The Dickey–Fuller test confirmed that the series became stationary after second differencing. Based on the analysis of autocorrelation and partial autocorrelation functions, as well as the Akaike and Bayesian information criteria, the ARIMA(1,2,1) specification was selected as the most appropriate model. The findings demonstrate the growing economic importance of tourism based on natural tourism resources in Uzbekistan. At the same time, the projected increase in tourist flows implies a higher intensity of resource use, which makes rational management, stronger environmental monitoring, and scientifically grounded regional tourism planning increasingly necessary. The study substantiates tourist-flow forecasting as an important analytical instrument for assessing the efficiency of natural tourism resource use and for shaping sustainable tourism policy.

Keywords: Natural tourism resources · sustainable tourism · tourism demand forecasting · ARIMA model · resource efficiency

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 630–642, 2027.
https://doi.org/10.1007/978-3-032-30530-5_40

1 Introduction

Contemporary global development trends demonstrate that tourism has emerged as one of the fastest-growing sectors of the world economy. At the same time, the long-term sustainable growth of the tourism industry largely depends on the extent to which natural tourism resources are utilized in a rational and efficient manner. Natural landscapes, climatic conditions, water resources, mountainous areas, and protected natural sites constitute the principal sources of tourist attractiveness in many destinations. According to the World Tourism Organization (UN Tourism), in 2024 the tourism sector generated an additional 27.4 million new jobs, supporting approximately 348 million jobs worldwide; furthermore, 1.4 billion international tourist arrivals were recorded globally, while tourism export revenues reached USD 1.9 trillion [1]. In this context, tourism's contribution to gross domestic product (GDP) amounted to USD 11.1 trillion in 2024, representing 10% of global GDP [2].

At present, within the framework of sustainable tourism development, scholarly research aimed at improving the economic mechanisms for the efficient use of natural tourism resources is expanding steadily. These studies place particular emphasis on assessing the condition of such resources, determining the efficiency of their use for tourism purposes, conducting comparative analyses of the experiences of different regions, and modeling and forecasting possible scenarios for tourism development. This demonstrates that the significance of statistical analysis and economic–mathematical modeling methods is continuously increasing in the study of tourism based on natural resources [3, 4].

In Uzbekistan, the development of a green economy, the promotion of environmental sustainability, and the rational use of existing natural resources are being identified as key priorities of state policy. This priority further reinforces the need, within the tourism sector, to develop economic mechanisms that comply with environmental requirements, improve methods for their assessment and monitoring, and foster the establishment of sustainable tourism infrastructure in the regions. In particular, the scientifically grounded assessment of the scale of natural resource use in tourist areas, the determination of the efficiency of such use, and the forecasting of future development trends through econometric methods are of crucial importance for the effective management of sustainable tourism.

From this perspective, the purpose of the present study is to reveal the economic essence of the efficiency of natural tourism resource use in Uzbekistan, to forecast the dynamics of the total number of tourists on the basis of the ARIMA model, and, based on the results obtained, to develop practical proposals and recommendations for the advancement of sustainable tourism. To achieve this objective, theoretical and methodological approaches to natural tourism resources are analyzed, the temporal dynamics of tourist flows are examined, an econometric forecasting model is applied, and scientific as well as practical conclusions are formulated on the basis of the findings. Such an approach makes it possible to provide a deeper assessment of the economic efficiency of natural tourism resource use, improve the quality of strategic decision-making in the tourism sector, and ensure sustainable tourism development at the regional level [5].

2 Literature Review

The sustainable development of the tourism sector largely depends on the availability of existing resources, particularly natural tourism resources. In many regions, unique natural landscapes, mountain air, therapeutic water sources, and protected natural sites constitute the principal factors attracting tourist flows. For this reason, natural tourism resources are regarded not only as the resource base of tourism, but also as a factor stimulating regional economic activity.

In the scholarly literature, the concept of "natural tourism resources" has been interpreted in different ways. In particular, Horun Ar Rashid defines these resources as natural factors that attract tourists to a specific area and includes within their composition geomorphological, climatic, hydrographic, and biogeographical resources, as well as protected natural heritage sites [6]. Other approaches interpret natural tourism resources as a set of natural factors that are not created by humans, exist in nature, and are suitable for use in tourism activities. In the present study, this concept is understood as a system of natural factors that shape tourist attractiveness and exert a direct or indirect influence on tourist flows [7]. P.K. Thornton and co-authors emphasize that problems related to the management of natural tourism resources and the environment are particularly pronounced in low-income countries [8]. In their view, this situation arises from the interrelationship of economic, social, and institutional problems. Natural resources constitute an important source of livelihood for such countries, and their improper management may lead to ecological imbalance and pose a threat to long-term sustainability. This approach is also highly relevant in the context of Uzbekistan, since many forms of tourism are directly based on the use of natural resources [9]. D.A. Ruban defines natural landscape elements that are considered valuable for tourism and attractive to visitors as tourism-related natural resources. He identifies geomorphological, geological, hydrological, and biological resources as the main components of this category [10]. This approach makes it possible to assess natural tourism resources not merely as aesthetic or recreational objects, but also as the material foundation of tourism development.

The analyzed studies primarily focus on the territorial distribution of natural resources, the factors enhancing tourist attractiveness, the innovative development of the tourism sector, the improvement of infrastructure, and issues of economic management [11]. At the same time, certain attention has also been paid to the influence of natural resources on tourism potential, the economic aspects of their sustainable use, and the territorial organization of tourism. It should be particularly emphasized that the existing studies provide an important scientific and methodological foundation for the development of the sector. However, in the context of Uzbekistan, the issue of comprehensively assessing the economic mechanism for the efficient use of natural tourism resources in relation to the future dynamics of tourist flows has not been sufficiently explored empirically [12, 13]. In particular, limited attention has been given to approaches that treat the total number of tourists as an integral indicator of the effectiveness of natural resource use, forecast it on the basis of time-series analysis, and, on this basis, assess the potential pressure exerted on resources.

3 Materials and Methods

This study employed annual time-series data covering the period from 2010 to 2024. The total number of tourists in Uzbekistan was selected as the main variable of analysis. This indicator was calculated as the sum of domestic tourists and inbound foreign tourists. Such an approach provides a more comprehensive assessment of the country's tourism potential, as tourism development is determined by the combined effects of both domestic and international tourist activity.

To evaluate the future dynamics of tourist flows, the AutoRegressive Integrated Moving Average (ARIMA) model was applied. This model is widely used in time-series forecasting because it captures the internal structure of a variable through its trend, autocorrelation pattern, and random disturbances. In the present study, the ARIMA approach was considered appropriate because the objective was to forecast the total number of tourists on the basis of its historical dynamics without explicitly incorporating exogenous explanatory variables. In this respect, the model makes it possible to identify and project the underlying temporal movement of tourist flows on the basis of their own stochastic structure.

In general form, the ARIMA model can be expressed as *ARIMA*(*p,d,q*), where p denotes the autoregressive order, d the degree of differencing, and q the moving average order. The general specification of the model is given as follows:

$$\Delta^2 Y_t = c + \sum_{i=1}^{p} \phi_i \Delta^2 Y_{t-1} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \quad (1)$$

where Y_t -is the total number of tourists in period tc is a constant term,

ϕ_i represents autoregressive coefficients, θ_j denotes moving average coefficients, and ε_t is the random error term. This approach allows modeling the stochastic nature of tourist flows while accounting for the serial dependence of the data series.

Theoretically, the ARIMA methodology is based on the Box–Jenkins approach, which consists of three main components: autoregression (AR), integration (I), and moving average (MA). The autoregression element expresses the dependence of the current value on its previous values. The integration part eliminates the non-stationarity of the series by differentiation. The moving average component reflects the influence of random errors in previous periods on the current value. It is precisely because of these features that the ARIMA model is particularly convenient for forecasting tourism indicators characterized by stable continuity, structural fluctuations, and non-stationary time-series behavior.

Before evaluating the model, the stationarity of the series was checked using the Dickey–Fuller test. The analysis showed that the total number of tourists was initially non-stationary: the test statistic was -0.254, and the p-value was 0.9318. The series did not reach a stationary state even after first-order differentiation, with the test statistic equal to -2.444 and the p-value equal to 0.1296. However, after second-order differentiation, the series reached a stationary state. This is confirmed by the test statistic being -3.108 and the p-value being 0.0259. Therefore, the integration level in the model was assumed to be $d = 2$. The second-order differentiation process is expressed as follows:

$$\Delta^2 Y_t = Y_t - 2Y_{t-1} + Y_{t-2} \quad (2)$$

This transformation was necessary to eliminate the trend component and stabilize the mean of the series prior to forecasting.

After the series was stationary, the autoregressive and moving-average orders of the model were determined from the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of the second-order differentiated series. The ACF plot showed significant peaks at the first and second lags, suggesting that q = 1 or q = 2 might be appropriate. Similarly, the PACF plot showed significant partial autocorrelations at the first and second lags, suggesting that p = 1 or p = 2 might be appropriate. On this basis, several alternative ARIMA models were evaluated. In this case, different combinations of model parameters were tested, and values of 0, 1, and 2 were obtained for the indicators p and q.

The final model was selected using the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), which are commonly employed to identify the most parsimonious specification with the best statistical fit. Among the estimated alternatives, the ARIMA(1,2,1) model produced the lowest information criterion values (*AIC = 268.241; BIC = 270.50*) and was therefore selected as the optimal forecasting model. The estimated form of the selected specification can be written as:

$$\Delta^2 Y_t = c + \phi_1 \Delta^2 Y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t \tag{3}$$

The estimated coefficients of the selected model further confirmed its empirical adequacy. In particular, the constant term was estimated at 300.479 ($p = 0.4939$), the autoregressive coefficient θ_1 at 0.317733 ($p = 0.2491$), and the moving average coefficient θ_1 at −0.999998 ($p < 0.0001$). The statistical significance of the moving average term indicates that short-term shocks play an important role in shaping changes in the differenced tourist series. In addition, the model exhibited acceptable goodness-of-fit characteristics, including R2 = 0.664833, adjusted R2 = 0.634364, and a log-likelihood value of −130.1206.

To assess the adequacy of the selected model, additional diagnostic tests were performed on the residual terms. First, the ARCH test was used to determine whether the residuals were heteroskedastic. According to the test result, the statistical indicator was 0.384078, and the corresponding p-value was 0.535429. Since this value was higher than the 5% significance level, the null hypothesis of the absence of the ARCH effect was not rejected. This means that the residual variance remained stable over time.

Then, the Breusch–Godfrey LM test was performed to determine the presence of autocorrelation in the residuals up to the third lag. The test statistic was 5.23675, indicating no statistically significant autocorrelation in the residuals. In general, these diagnostic tests confirm that the ARIMA(1,2,1) model is sufficiently suitable from an econometric perspective and for forecasting. Based on the evaluated ARIMA(1,2,1) model, forecast values for the total number of tourists in Uzbekistan for 2025–2030 were calculated. During the forecasting process, not only expected values, but also standard errors and 95% confidence intervals were determined. This allowed us to assess the likely future direction of tourist flows and indicate the level of uncertainty associated with the forecast results. Model evaluation, selection of the most optimal option, diagnostic tests and forecast calculations were all carried out using the Gretl statistical program.

The forecast results allow us to empirically assess the expected growth in tourist flows in Uzbekistan and anticipate potential pressure on natural tourist resources. In

this regard, this model is not only a forecasting tool but also an important analytical instrument for developing strategic decisions on sustainable tourism development, the rational use of resources, and the long-term management of tourism growth.

4 Results

4.1 Dynamics of Tourist Flows in Uzbekistan

The dynamics of tourist flows in Uzbekistan during 2010–2024 showed a steady growth trend in both domestic and foreign tourism (Fig. 1). In 2010, the number of domestic tourists was 2,065 thousand, and the number of foreign tourists visiting the country was 741 thousand. During the study period, domestic tourism, in particular, developed at a relatively rapid pace. In particular, by 2013, the number of domestic tourists reached 6,564 thousand, indicating the growing role of domestic tourism in the country's tourism development.

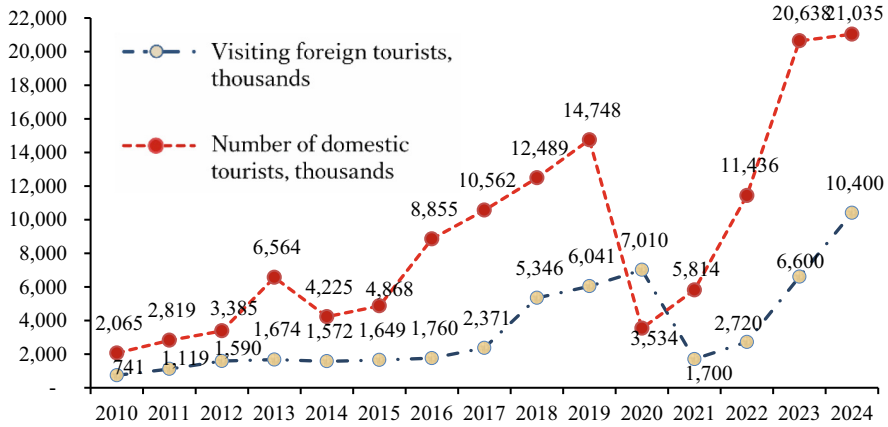


Fig. 1. Dynamics of domestic and foreign tourist flows in Uzbekistan, 2010–2024.

In particular, a significant increase was observed in 2016–2019. During this period, natural resource-based tourism, including ecological, mountain, and agrotourism, developed more actively. Despite a temporary decline during the pandemic years, demand for tourism related to natural areas remained stable and entered a clear recovery phase in 2021–2024. By 2024, the number of domestic tourists exceeded 21 million, while the number of foreign tourists surpassed 10 million. These patterns indicate that tourist activity in Uzbekistan has regained momentum and continues to expand in the post-pandemic period.

4.2 Stationarity of the Tourist Series

Before estimating the forecasting model, the stationarity of the total number of tourists was examined using the Dickey–Fuller test [14, 15]. The results are presented in Table 1.

As shown, the tourist series was non-stationary in its level form, with a test statistic of −0.254 and a p-value of 0.9318. The first-differenced series also remained non-stationary, as indicated by a test statistic of −2.444 and a *p*-value of 0.1296. However, after second differencing, the series became stationary, with a test statistic of −3.108 and a p-value of 0.0259. These results justify the specification of the integration order as $d = 2d$ in the ARIMA framework.

Table 1. Dickey–Fuller test results for the stationarity of the total number of tourists

Variable	Test statistic	Critical value (5%)	p-value	Conclusion
TurSoni	−0.254	−3.000	0.9318	Non-stationary, I(0)
d_TurSoni	−2.444	−3.000	0.1296	Non-stationary, I(1)
d2_TurSoni	−3.108	−3.000	0.0259	Stationary, I(2)

Following the results of the stationarity test, the tourist series was transformed through second differencing, and the forecasting model was subsequently estimated on the stationary series.

4.3 Identification and Selection of the ARIMA Model

The identification of the autoregressive and moving average orders was based on the analysis of the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the second-differenced tourist series (Figs. 2 and 3). The ACF plot exhibited notable spikes at the first and second lags, suggesting plausible values of q = 1or q = 2q. Similarly, the PACF plot displayed significant spikes at the first and second lags, indicating that p = 1p 2p = 2 could be considered appropriate. On this basis, several candidate ARIMA(p,2,q) specifications were estimated in the Gretl statistical software package, where $p \in \{0, 1, 2\}$ and $q \in \{0, 1, 2\}$. Model selection was carried out using the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). Among the alternative specifications, the ARIMA(1,2,1) model yielded the lowest information criterion values (AIC = 268.241; BIC = 270.50), and was therefore selected as the optimal model for forecasting the total number of tourists.

4.4 Estimation Results of the ARIMA(1,2,1) Model

The estimation results of the selected ARIMA(1,2,1) model are reported in Table 2. The coefficient of the moving average term ($\theta_1 = -0.999998$) was found to be highly statistically significant ($p < 0.0001$), while the constant and autoregressive coefficients were not statistically significant at conventional levels. This suggests that short-term shocks and error-adjustment effects play an important role in explaining the dynamics of the differenced tourist series.

The model also demonstrated satisfactory statistical properties. In particular, the coefficient of determination reached $R2 = 0.664833$, while the adjusted R^2 amounted

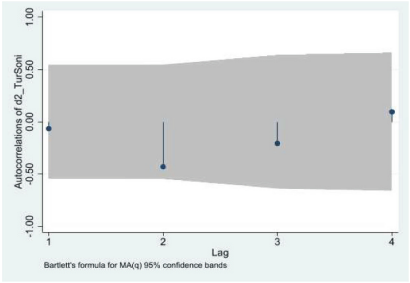


Fig. 2. Autocorrelation function (ACF) of the second-differenced series of the total number of tourists

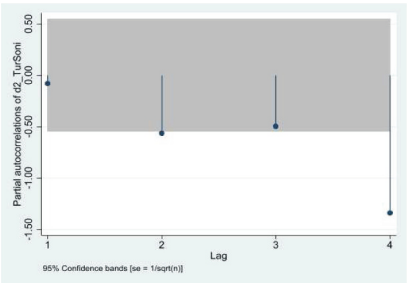


Fig. 3. Partial autocorrelation function (PACF) of the second-differenced series of the total number of tourists

to 0.634364. The log-likelihood value was estimated at -130.1206. These results indicate that the selected model captures a substantial proportion of the variation in the transformed series and is suitable for forecasting purposes.

Table 2. Estimation results of the ARIMA(1,2,1) model for the total number of tourists

	Coefficient	*Std. Error*	*z*	*p-value*	
const	300.479	439.179	0.6842	0.4939	
phi_1	0.317733	0.275709	1.152	0.2491	
theta_1	−0.999998	0.204465	−4.891	<0.0001	***
Mean dependent var	235.7769		S.D. dependent var	6532.384	
Mean of innovations	−228.2626		S.D. of innovations	4972.484	
R-squared	0.664833		Adjusted R-squared	0.634364	
Log-likelihood	−130.1206		Akaike criterion	268.2411	
Schwarz criterion	270.5009		Hannan-Quinn	267.7767	

4.5 Residual Diagnostic Results

To evaluate the adequacy of the estimated ARIMA(1,2,1) model, additional residual diagnostic tests were conducted. First, an ARCH test of order one was applied to detect possible heteroskedasticity. The test statistic was 0.384078, with a corresponding p-value of 0.535429. Since this value exceeds the 5% significance level, the null hypothesis of no ARCH effects cannot be rejected, indicating that the residual variance remains stable over time.

Second, the Breusch–Godfrey LM test was performed to assess the presence of residual autocorrelation up to the third lag. The test statistic amounted to 5.23675. On this basis, no substantial residual autocorrelation was identified in the model residuals. Taken together, the diagnostic results confirm that the selected ARIMA(1,2,1) specification is econometrically adequate and statistically reliable for forecasting the future dynamics of tourist flows (Table 3).

Table 3. Residual diagnostic test results for the ARIMA(1,2,1) model

Diagnostic test	Test statistic	p-value	Conclusion
ARCH test (lag 1)	0.384078	0.535429	No heteroskedasticity detected
Breusch–Godfrey LM test (up to lag 3)	5.23675	n/a	No substantial residual autocorrelation detected

4.6 Forecast Results for 2025–2030

Based on the estimated ARIMA(1,2,1) model, forecast values for the total number of tourists in Uzbekistan for 2025–2030 were calculated. The forecast results, standard errors and 95% confidence intervals are presented in Table 4, while the forecast direction is depicted in Fig. 4. The results obtained indicate that a stable growth trend in tourist flows will be maintained throughout the forecast period.

More specifically, the total number of tourists in Uzbekistan is expected to reach 35,922.4 thousand in 2025. The 95% confidence interval for this indicator is from 26,176.5 thousand to 45,668.3 thousand. The growth trend will continue in the following years. In particular, by 2030, the forecast value is expected to reach 62,843.8 thousand people, with a 95% confidence interval of 30,326.8 thousand to 95,360.9 thousand people. In subsequent years, the expansion of the confidence interval indicates that uncertainty increases as the forecast period lengthens. This is typical for medium-term time series forecasts.

The indicators of forecast quality also confirm the model's practical usefulness. In particular, the Mean Error was −228.26, the Root Mean Squared Error (RMSE) was 4972.5, and the Mean Absolute Error (MAE) was 3437.2. Also, the Mean Absolute Percentage Error (MAPE) was 29.221, and the Theil's U statistic was 0.7946. These results show that the ARIMA(1,2,1) model is a sufficiently robust and acceptable tool for forecasting the total number of tourists in Uzbekistan in the medium term.

Table 4. Forecast values of the total number of tourists in Uzbekistan for 2026–2030

Year	Forecast value	Standard error	95% confidence interval
2026	40706.8	8225.55	24585.0–56828.6
2027	45790.7	10836.3	24552.0–67029.4
2028	51174.7	13017.9	25660.1–76689.3
2029	56859.0	14907.7	27640.4–86077.6
2030	62843.8	16590.6	30326.8–95360.9

Note: *Mean Error = -228.26; Root Mean Squared Error (RMSE) = 4972.5; Mean Absolute Error (MAE) = 3437.2; Mean Percentage Error (MPE) = -8.6905; Mean Absolute Percentage Error (MAPE) = 29.221; Theil's UUU statistic = 0.7946; Bias proportion (UM) = 0.0021; Regression proportion (UR) = 0.1181; Disturbance proportion (UD) = 0.8798.*

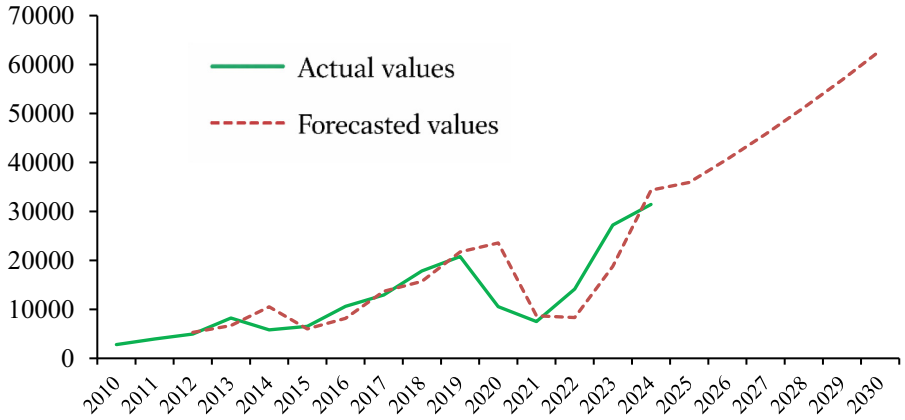


Fig. 4. Forecast trajectory of the total number of tourists based on the ARIMA(1,2,1) model.

Overall, the forecast results indicate that the total number of tourists in Uzbekistan will continue to grow steadily in the medium term. The most reliable forecast intervals fall in the near future, especially 2026–2027. The range of uncertainty increases as the forecast period lengthens. Nevertheless, the central forecast indicators indicate that tourist flows will continue to grow consistently in all years.

5 Discussion

The study's results show that the total number of tourists in Uzbekistan will continue to grow over the forecast period. This means that the demand for tourism in the country will remain high in the medium term. In the context of this study, this growth not only reflects the expansion of the tourism sector but also indicates the growing economic importance of natural tourist resources. Since a significant part of tourist activity in Uzbekistan is directly related to mountainous areas, protected natural areas, ecological destinations and other nature-based tourism objects [16].

The results obtained are consistent with previous scientific studies that emphasize that natural resources are among the main factors in tourism development [17, 18]. In previous empirical stages of the study, it was found that indicators directly related to natural resources, particularly protected areas and environmental sustainability factors, have a positive, statistically significant impact on tourist flows [8, 19–21]. In this regard, this forecast based on the ARIMA model does not contradict previous conclusions, but rather complements them. That is, it indicates that the growth trend in the number of tourists is likely to continue in the coming years.

Another important aspect of the study is that the projected increase in tourist flows has a dual meaning. On the one hand, this situation creates opportunities to increase tourism revenues, expand employment, and strengthen regional economic activity [22]. On the other hand, it means that the intensity of use of natural tourist resources may also increase. This is especially important for regions with limited ecological capacity. Since in such places, an increase in tourist numbers can put greater pressure on natural landscapes and recreational infrastructure. Therefore, the projected growth in the number

of tourists should be interpreted not only as a positive economic indicator but also as a signal of the need for more balanced resource management.

The fact that the selected ARIMA(1,2,1) model provides statistically acceptable results further strengthens the analytical significance of the forecast. The absence of significant heteroscedasticity and significant autocorrelation in the model residuals confirms its reliability as a basis for medium-term forecasting. At the same time, the widening of the confidence interval in the later years of the forecast period indicates an increase in uncertainty over time. This is a natural setting for time-series forecasting. Nevertheless, the forecast remains stable, indicating that tourist activity will continue to grow in the future.

Another important aspect is the clear reflection of the post-pandemic recovery process in the time series. After the decline observed during the pandemic, the flow of tourists has begun to recover and has entered a phase of stable growth. This indicates that the demand for tourism in Uzbekistan, including interest in natural and recreational destinations, is relatively stable. Therefore, the forecast results can be viewed as additional evidence that natural resource-based tourism will continue to play an important role in the country's long-term tourism development.

Overall, the results of the study suggest that forecasting the total number of tourists can serve as an important analytical tool for assessing the efficiency of future use of natural tourist resources [18, 23]. In this sense, the expected growth in tourist flows provides an important empirical basis for improving the tourism planning system, strengthening environmental control, and ensuring a more sustainable balance between tourism expansion and natural resource conservation.

6 Conclusion

This study shows the need to assess the effective use of natural tourist resources in Uzbekistan not only as an environmental or territorial issue, but also as an important economic factor directly affecting tourism development. The forecast results from the ARIMA(1,2,1) model confirm that the total number of tourists will steadily increase in the medium term. This means that demand for tourism in the country will continue to expand, and the role of natural resource-based tourism in the national tourism system will further strengthen.

One of the important conclusions of the study is that the growth of tourist flows simultaneously creates economic opportunities and exacerbates resource management problems. On the one hand, an increase in the number of tourists contributes to higher tourism revenues, expanded employment, and increased regional economic activity. On the other hand, this process further increases the pressure on natural tourist resources, especially in ecologically sensitive and attractive regions. Therefore, tourism development policy should not be limited to economic priorities, but should also be based on the principles of environmental sustainability and the rational use of resources.

The results obtained show that forecasting tourist flows is not a simple auxiliary analytical method, but an important practical tool for effective management of sustainable tourism. Assessments based on forecast results provide a solid empirical basis for determining future tourist pressure, improving the efficiency of natural resource use,

enhancing regional tourism strategies, and strengthening the environmental monitoring system. In this sense, the research results provide a scientific and practical basis for ensuring a balanced approach between tourism growth and the preservation of natural tourist resources in Uzbekistan.

Foundation. This work was supported by the Innovation Grant "Methodological Foundations for Organizing Inclusive Tourism (Based on the Example of Khatirchi, Nurata, and Kiziltepa Districts of Navoi Region) [Grant No. IL-9124093892].

References

1. UN Tourism World Tourism Barometer | Global Tourism Statistics. https://www.untourism.int/un-tourism-world-tourism-barometer-data. Accessed 13 Apr 2026
2. Travel & Tourism Economic Impact Research (EIR). https://wttc.org/research/economic-impact. Accessed 13 Apr 2026
3. Farrukh, B., Younis, I., Longsheng, C.: The impact of natural resource management, innovation, and tourism development on environmental sustainability in low-income countries. Resour. Policy **86**, 104088 (2023). https://doi.org/10.1016/j.resourpol.2023.104088
4. Safarov, B., Mansurova, N., Hasanov, H., Samiyev, S.: Strategic development of tourism and hospitality for sustainability in Uzbekistan. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15899, pp. 187–196. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97663-6_16
5. Mukhammedov, M., Safarov, B., Hasanov, H., Isxakova, S., Buzrukova, M., Hassan, T.H.: Methods for optimizing property relations as a key aspect of stimulating the development of the tourism industry. Geoj. Tour. Geosites **58**, 484–491 (2025). https://doi.org/10.30892/gtg.58145-1430
6. Rashid, M.H.A.: Types of Tourism Resources. https://limbd.org/types-of-tourism-resources/. Accessed 13 Apr 2026
7. Gjorgievski, M., Kozuharov, S., Nakovski, D.: Typology of recreational-tourism resources as an important element of the tourist offer. UTMS J. Econ. **4**, 53–60 (2013)
8. Thornton, P.K., Van De Steeg, J., Notenbaert, A., Herrero, M.: The impacts of climate change on livestock and livestock systems in developing countries: a review of what we know and what we need to know. Agric. Syst. **101**, 113–127 (2009). https://doi.org/10.1016/j.agsy.2009.05.002
9. Safarov, B., et al.: Prospects of agrotourism development in the region. Economies **12** (2024). https://doi.org/10.3390/economies12120321
10. Ruban, D.A.: Natural resources of tourism: towards sustainable exploitation on a regional scale. Sustainability **13**, 6685 (2021). https://doi.org/10.3390/su13126685
11. Safarov, B., Mansurova, N., Hasanov, H., Xayrullayev, A., Sanayeva, R., Yu, J.: Analysis of factors influencing the development of health and wellness tourism in the context of sustainable urban and regional development in Uzbekistan. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15890, pp. 405–414. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97606-3_27
12. Box, G.E., Jenkins, G.M., Reinsel, G.C., Ljung, G.M.: Time Series Analysis: Forecasting and Control. Wiley, Hoboken (2015)
13. Tuxtabekov, K., Buzrukova, M., Hasanov, H., Salnikova, E.: Forecasting tourism services exports for sustainable development: a case study of Uzbekistan. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15899, pp. 176–186. Springer, Cham (2026). https://doi.org/10.1007/978-3-031-97663-6_15

14. Porto, N., Rucci, A.C., Ciaschi, M.: Tourism accessibility competitiveness. A regional approach for Latin American countries. Investigaciones Regionales J. Reg. Res., 75–91 (2018)
15. Natalia, P., Clara, R.A., Simon, D., Noelia, G., Barbara, A.: Critical elements in accessible tourism for destination competitiveness and comparison: principal component analysis from Oceania and South America. Tour. Manag. **75**, 169–185 (2019). https://doi.org/10.1016/j.tourman.2019.04.012
16. Arabov, N., et al.: Shaping the future of Uzbekistan's tourism: an in-depth analysis of infrastructure influence and strategic planning. JEECAR **11**, 53–65 (2024). https://doi.org/10.15549/jeecar.v11i1.1478
17. Shimizu, T., Okamoto, H.: Tourism infrastructure and the environment: how does environmental regulation affect welfare, tourism industry, and domestic wage inequality? JER **75**, 147–179 (2024). https://doi.org/10.1007/s42973-021-00109-4
18. Grenier, A.A.: The nature of nature tourism. University of Lapland, Faculty of Social Sciences (2004)
19. Ilieș, D.C., et al.: Preserving Cultural heritage: enhancing limestone durability with nano-TiO2 coating. Heritage **7**, 4914–4932 (2024). https://doi.org/10.3390/heritage7090232
20. Caciora, T., et al.: Comprehensive analysis of classroom microclimate in context to health-related national and international indoor air quality standards. Front. Public Health **12**, 1440376 (2024). https://doi.org/10.3389/fpubh.2024.1440376
21. Caciora, T., et al.: Advanced semi-automatic approach for identifying damaged surfaces in cultural heritage sites: integrating UAVs, photogrammetry, and 3D data analysis. Remote Sens. **16**, 3061 (2024). https://doi.org/10.3390/rs16163061
22. Amirov, A., Nurimbetov, R., Arabov, N., Boltaev, B., Hasanov, H., Nasimov, I.: Exploring the dynamics of human capital and economic growth in central Asian countries: an empirical analysis using panel data. Glob. Bus. Rev., 09721509251415274 (2026). https://doi.org/10.1177/09721509251415274
23. Anthony, E.J.: The status of beaches and shoreline development options on the French Riviera: a perspective and a prognosis. J. Coast. Conserv. **3**, 169–178 (1997). https://doi.org/10.1007/BF02905242

Author Index

A
Acu, Ana Maria 124
Alecsa, Cristian Daniel 124
Annunziata, Alfonso 22, 39
Anselmi, Chiara 3
Aversano, Lerina 351

B
Belcore, Orlando M. 527
Bogdanov, Alexander 481
Burciu, Stefan 585

C
Cardone, Angelamaria 184, 199
Catabriga, Lucia 313
Chernigov, Aleksei 493
Ciaramella, Angelo 173
Cocco, Eugenio Maria 95
Codroiu, Denis 585
Coelho, C. 255, 285
Coelho, Cecília 272
Conte, Dajana 199
Costa e Silva, Eliana 241
Costa, Fernanda 241
Costa, M. Fernanda P. 272, 285
Cutini, Valerio 60

D
Danese, Maria 371, 385
De Luca, Pasquale 173
De Lucia, Carmine 221
Deri, Federica 3, 60
Di Gangi, Massimo 527
Di Nardo, Emanuel 173
Di Pascale, Gerardo 184
Dinu, Alexandru 298
Dinu, Oana 585
Díaz de Alba, Patricia 199
Do, Huy Minh Dung 333
Donatelli, Marco 108, 156
Dorobantu, Iulia 585

F
Fazia, Celestina 615
Ferrás, L. L. 272, 285
Frunzete, Mădălin 298

G
Galantucci, Stefano 351
Garzia, Emilio 173
Gerace, Ivan 77
Gomes, André Paz 313

H
Ha, Muon 403
Ha, Thien Hai 333

I
Ibatullin, Evgeniy 481
Iscaro, Samira 199

K
Koloskov, Victor 464
Korkhov, Vladimir 447, 493
Kukharev, G. A. 415

L
Lasaponara, Rosa 385
Leonessa, Vita 142

M
Manh, Tran Duc 403
Mara, Federico 60
Marcellino, Livia 173
Mastellone, Andrea 221
Mastroianni, Michele 351
Matrone, Alfonso 221
Maurino, Pietro 108

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16763, pp. 643–644, 2027.
https://doi.org/10.1007/978-3-032-30530-5

Modan, Sara-Gabriela 298
Morgado, M. L. 285
Moschella, Marialuisa 556, 570
Murgante, Beniamino 22, 39
Musolino, Giuseppe 543, 556, 570

N
Nargiza, Mansurova 630
Nguyen, Xuan -Hien 403
Nicoloso, Sara 95
Niggemann, O. 255
Nurjemal, Kakabayeva 433

P
Pagano, Giovanni 199
Panuccio, Paola 543
Paternoster, Beatrice 199
Perciante, Felice 385
Pershin, Anton 464
Pezzica, Camilla 3
Pham, Duc Hoang 333
Pham-Quoc, Cuong 333
Phan, Thanh-Son 403
Polimeni, Antonio 527
Porcelli, Andrea 351

R
Rebelo, M. 285
Rindone, Corrado 543, 598
Roman, Alina 585
Rusca, Aura 585
Rusca, Florin 585
Russo, Antonio 598
Russo, Francesco 543, 556, 570

S
Sabia, Canio Alfieri 371
Safarov, Bahodirhon 630
Safonenko, Aleksandr 447
Santos, Isaac P. 313
Sapio, Raffaella 221
Scarlato, Giuseppe 156
Sgro, Domenico 556, 570
Shchegolev, A. D. 415
Shchegoleva, N. L. 415
Shchegoleva, Nadezhda 433
Shichkina, Yulia 403
Silva, Maria 241
Stankova, Elena 511
Subbotin, Alexey 511
Szylagyine Czimre, Klara 615

T
Tătar, Corina-Florina 615
Toraldo, Gerardo 221
Toshkhujaev, Mustafokhon 630
Travaglini, Arianna 124, 142

V
Villar, E. 272
Vinti, Gianluca 124, 142
Vitetta, Antonino 543
Vicari Aversa, Clara Stella 615

Z
Zhukova, Nataly 511
Zullo, Ferdinando 221