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

LNCS 16766

Computational Science and Its Applications – ICCSA 2026 Workshops

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

9 Part IX



Springer

Lecture Notes in Computer Science **16766**

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 IX

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 978-3-032-30538-1 ISBN 978-3-032-30539-8 (eBook)
https://doi.org/10.1007/978-3-032-30539-8

© 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

Virtual Reality, Augmented Reality and Applications (VRA 2026)

Indoor Pilot 3D Representation of Comfort Conditions in Mixed Reality Environment 3
Stelios Kouzeleas, Ioannis Tsolakidis, and Eleni Vlachonasiou

Smart City Digital Twin Framework for Semantic Indexing and Dynamic Modelling of 3D Buildings 21
Marco Fanfani, Federico Galli, Chiara Garau, Paolo Nesi, and Gianni Pantaleo

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

Integrating GIS-Based Spatial Analysis and Structural Modeling for the Adaptive Reuse of the Santa Chiara Dam 37
Enrica Vecchi and Marco Zucca

Integrated Geophysical Study at the Colonel Henry Anderson Morshead Burial Memorial, Valletta: Preliminary Results 54
Sebastiano D'Amico, Luca Piroddi, Emanuele Colica, Luciano Galone, and Mario Farrugia

Investigating the Castello Archaeological Park (Cersosimo, Italy) Using a Combination of Geophysical Prospections, Aerial and Terrestrial Photogrammetry 68
Marilena Cozzolino, Vincenzo Gentile, and Paolo Mauriello

Assessment of Low-Cost GNSS Receiver Performance in Urban Environments 86
Giuseppina Vacca, Enrica Vecchi, and Antonio Pala

Integration of 3D Ultrasonic Tomography, Photogrammetry, and Mixed Reality for the Structural Assessment of the Marble Statue "Caino" 104
Marco Vincenzo Majani, Alessandra Carollo, Raffaele Martorana, and Patrizia Capizzi

PHD Showcase Papers

An Efficient and Stable Solver for the Shallow Water Equations on Regular Latitude-Longitude Grids Using Spherical Fourier Transforms 117
Suzune Nomura and Takeshi Enomoto

Explainable and Accurate Valvular Heart Disease Screening Using Spectrogram-Based 2D-CNNs on PCG Signals 135
Talha Ahmed Siddiqui, Sania Akhtar, Muhammad Taha Mansoor, Badee Baltow, Sadaq-Ur Rehman, and Noor Ul Huda

Revisiting Ray-Tracing Algorithm for Future Limb-Scanning Missions 153
Francesco Pio De Cosmo and Luca Sgheri

Short Papers

Artificial Intelligence for Photovoltaic Power Forecasting in Energy Management 173
Guillermo Blanco, Marco Melgarejo, Alejandro Pérez, and Camilo Palazuelos

Economic Assessment of Inclusive Tourism in a Silk Road Heritage Destination: A Monte Carlo Simulation Study of Samarkand, Uzbekistan 186
Bahodirhon Safarov, Habibullo Hasanov, Makhina Buzrukova, and Elena Salnikova

On Battery Storage Management by Sparse Linear Programming 197
Ana Maria A. C. Rocha, Pablo Guerrero-García, and Eligius M. T. Hendrix

On Subproblem Solving in an Augmented Lagrangian Algorithm for LP with Upper Bounds 206
Eligius M. T. Hendrix, Pablo Guerrero-García, and Ana Maria A. C. Rocha

Resilient Mobility Under Multiple Shocks: The SCUDO Project Experience . . . 217
Francesco Bruzzone, Matilde Gasparetto, and Silvio Nocera

ViBERT-ESG: Advancing ESG Classification in Vietnamese with ViEnESG-100 Dataset 226
Vanmanh Nguyen, Ductrung Tran, Kienlam Do, Thiloan Bui, and Vanha Tran

Machine Learning Assisted Two-Level Mixed Finite Element Method for the Darcy–Forchheimer Model 236
Johnson Daddy Audu and Faisal Fairag

On the Travelling Wave Solutions of a Non-linear Reaction-Diffusion-Advection Model for Forest Fire Spread 247
Donato Pera, Cordula Reisch, Federica Di Michele, and Bruno Rubino

Anomaly Detection and Root Cause Analysis of Retainability Contextual Anomalies in Radio Access Networks 259
David Santos, Paula Amaral, João Martins, and Rui Ramos

THEMATIC2GREEN: Strengthening Innovation Capacity for a Place-Based Green Transition in Fragile ADRION Territories 270
Michele Greco, Francesco Scorza, Biagio Perretti, Francesco Paolo Del Giudice, Alfonso Annunziata, and Beniamino Murgante

A Unified Energy Decomposition and Delocalization Approach to Rare-Gas Excimer Bonding 281
Marcos Mandado

Molecular Dynamics Simulations of Nanoemulsions: Effect of the Initial Geometry and Concentration 291
Jose Manuel Hermida-Ramon and Sonia Losada-Barreiro

Transport Disruption Vulnerability in Europe: Who Is Worse Off? 300
Francesco Bruzzone, Hamza Begic, Silvio Nocera, and Constantinos Antoniou

Backlog Traps, Congestion and Sustainable Transition in Port Systems: A System Dynamics Approach 311
Ani Simonyan and Giuseppe Borruso

On the Behavior of PATE Under Class Imbalance and Non-IID Data 327
Alexander Bogdanov, Valery Khvatov, Yueh Chiang, and Nadezhda Shchegoleva

Traffic Flow Analysis of Closely Spaced Roundabouts and Signalized Intersections 338
Mirna Klobučar, Aleksandra Deluka-Tibljaš, Sanja Šurdonja, and Irena Ištoka Otković

Evaluating the Impact of Traffic Control System Failures on Intersection Level of Service Using Microsimulation 351
Denis Codroiu, Eugen Rosca, Aura Rusca, Maja Skuric, Oana Dinu, Valentina Radu, and Armand Stere

Control Point Detection for Network Alignment Using Large Language Model 365
Müslüm Hacar and Özge Öztürk Hacar

Active Reflectors and Passive Corner Reflectors for Deformation Monitoring Using DInSAR Interferometry: Some Remarks from Sentinel-1 Imagery Tests 374
G. Luzi, P. Espín-López, and Q. Gao

Statistical Anomaly Detection Using LC-Curves: An Application to Geophysical Time Series 384
Fuad Aleskerov and Aleksei Kosychev

A Robotic GPR Pipeline for Automated Subsurface Analysis in Multiple Scenarios 393
Nuria Rodríguez Calderón, Berta Viñes Santamaria, David Caballero Flores, Tahzeeb Hussain, and Vega Perez-Gracia

Decarbonising Long-Distance Transport: Modal Shift from Short-Haul Aviation to High-Speed Rail for Sustainable Mobility 405
Oana Dinu, Eugen Rosca, Ilona Costea, Andrea Alonso Ramos, Aura Rusca, Anamaria Ilie, and Cristian Marin

Nanofluid Injection in Poroelastic Media: Possible Applications in Environmental Science 416
Rosella Sampalmieri, Federica Di Michele, Donato Pera, and Bruno Rubino

Author Index 427

Virtual Reality, Augmented Reality and Applications (VRA 2026)

Indoor Pilot 3D Representation of Comfort Conditions in Mixed Reality Environment

Stelios Kouzeleas[1(✉)], Ioannis Tsolakidis[2], and Eleni Vlachonasiou[3]

[1] Department of Interior Architecture, International Hellenic University, Serres, Greece
stelios_kouzeleas@yahoo.fr
[2] Natural Environment & Climate Change Agency, Serres, Greece
i.tsolakidis@necca.gov.gr
[3] Department of Civil Engineer, International Hellenic University, Serres, Greece
vlachonasiou@ihu.gr

Abstract. This theoretical description of interdisciplinary research combines different scientific fields, to present a spatial digital interactive 3D representation system of indoor spatial properties, such as unhealthy emissions and building's energy rates, in virtual or mixed reality environment in order to record and evaluate both the energy and health status inside a building in an understandable way for the average habitant. A method of a 3D grid of modern sensors records rates based on voxels / point clouds technology is proposed with also a final comparison and an interactive sensor values evaluation in terms of health and energy. This results a 3D digital thematic representation of "cloud" rates of energy and health status inside the building in a mixed reality environment with a superimposed rates evaluation understandable table. Finally, the representation system allows both the average habitant and the political leaders to evaluate energetically and hygienically an indoor environment, by sensing appropriate actions and creating a social health policy.

Keywords: Immersive spatial 3D Representation · Augmented Reality · Mixed Reality · Digital Design · Energy & Health Building Design · Energy conditions simulation

1 Introduction

During the last decade, many 3D representation systems are deployed for maintaining optimal comfort conditions with the aid of live streams of sensor data. The Internet of Things (IoT) has been applied in various platforms to enhance the quality of life [1], while IoT and other digital sensors provide real-time feedback to building occupants and this information will be of great value since indoor conditions have a huge influence on the quality of their well-being [2] by using simulations of spatial elements and descriptive forms [3] with interdisciplinary approaches. The self-aware building systems aim to acquire dynamic faithful knowledge of buildings and learn to control themselves adaptively and perform processes [4], (Fig. 1).

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 3–20, 2027.
https://doi.org/10.1007/978-3-032-30539-8_1

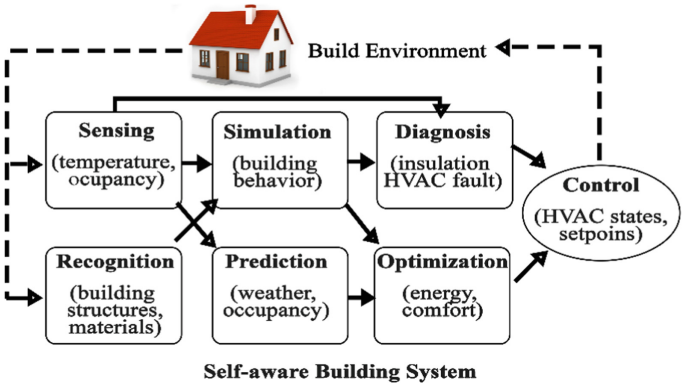


Fig. 1. A self-aware building system consists of several autonomous processes. Source: [5]

A specific need, among others, for well-being inside a building has led many studies, in recent years, on measurement and 3D representation of comfort conditions and / or important indoor spatial properties [2, 6]:

(1) building's energy design spatial properties, such as humidity, temperature, ventilation, thermal conditions, etc., and
(2) building's health design spatial properties, such as unhealthy emissions and pollutants, etc.

However, the combination of interdisciplinary techniques and technologies presents many advantages and limitations and must be adapted to the specific conditions and needs of each study.

2 Proposal Objectives and Challenges

The present research addresses the creation of an indoor pilot 3D representation of comfort conditions in a Mixed Reality environment with the use of interdisciplinary approaches and techniques. The proposed simulation system / methodology objectives are quite similar with previous studies of some of the same authors on nosocomial infections detection [7] and Air Quality emissions inside buildings [8] via biosensors particles spatial simulation in Mixed Reality environment. In this way, the proposed pilot 3D representation system will:

(1) describe in theoretical way a spatial modelling and simulation methodology of comfort conditions rates in order to contribute to a future implementation of a real-time 3D representation system,
(2) visualize temperature, humidity energy design and air pollution inside a building in 3D volumetric interactive 3d point cloud/voxel simulation with spatial reference inside either a virtual building or an AR building environment,
(3) test point cloud filtering algorithms for voxelization of a building,
(4) create a representative 3d models of objects in a minimum time by using scanning techniques,

(5) provide a framework for making a Post Occupancy Evaluation (POE) of energy efficiency (EE) in buildings' indoor environment with results visualization in 3d Virtual Environment.
(6) evaluate measurements rates in a simple graphical and ergonomic way for a non-expert user by comparing the measured rates with the optimal rates during the indoor walkthrough,
(7) contribute to create a wireless low-cost sensors grid system of comfort conditions measurement,
(8) define an adapted methodology of 3D effective indoor spatial modelling and mapping for MR with the use of low-cost system comfort conditions representation,
(9) contribute to understand the "problematic comfort areas" and implement a system by a statistic base knowledge to fix or improve these areas by sensing appropriate actions,
(10) visualize spatial phenomena with high degree of perception

3 State-of-the-Art and Innovation

3.1 Visualization Methods and Techniques

As this research is the extension of previous studies [7, 8], the state-of-the-art of this research proposal focuses on interdisciplinarity simulation & modelling approaches and techniques.

The Comfort conditions methods / techniques for volumetric visualization are based on several modern dynamic technologies, dealing, among others, with techniques in image overlay, OpenGL and VRML programming, special software kit [10], and with techniques in Wireless and telecommunication, VR, AR and MR technology, 3D spatial modelling [9], Building Information Modelling (BIM) integration, Voxel / points cloud representation, displays technology, energy visualization, etc.

VR-AR-MR techniques: While VR techniques immerse the user inside a synthetic environment, AR superimposes the virtual objects into the real world, supplementing reality instead of replacing it [11]. If the real world can be augmented with virtual objects, consequently the virtual world can be augmented with real scenes (views, objects). Such an environment is called Augmented Virtuality (AV), while Mixed Reality (MR) occupies the space between Real and Virtual Environment [11] (Fig. 2).

The basic techniques in AR [10], deal with (a) Tracking, (b) Result Visualization, (c) Interaction & collaboration and (d) Client-Server Network Architecture.

Wireless techniques: especially in AR environment face three fundamental issues, such as (a) mixing real and virtual environment, (b) tracking [13] and (c) wireless communication between users and backbone computers (Fig. 3).

Voxel modeling techniques: Voxelization or point clouds techniques refers to the segmentation of space into cubes of specific dimensions and specific structure by using a 3D cube table, called voxel [15] that can represent a wide range of 3D objects and their related properties [8, 12], such as material type, density, position, etc., as well as, environmental parameters such as the heat transfer in a room through the segmentation of space by using sensors and voxel modelling techniques [16].

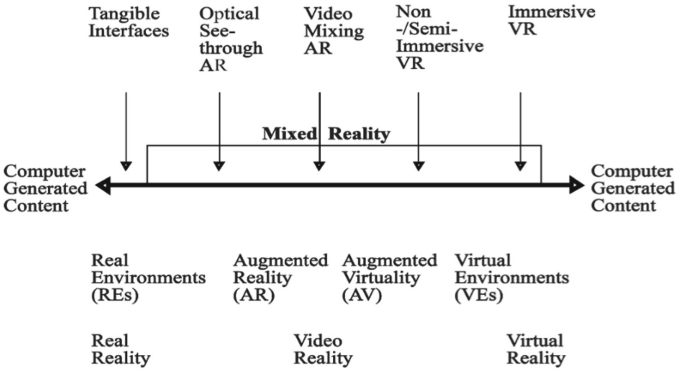


Fig. 2. Reality-Virtuality Continuum [11]

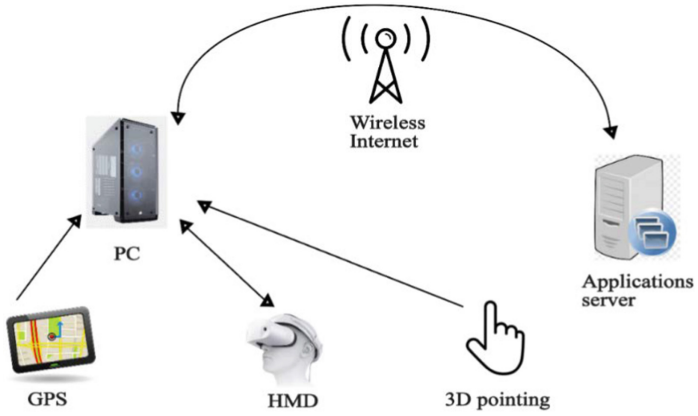


Fig. 3. Wireless AR application (M. Groves) [14]

Point clouds techniques to structure voxel grid space: The conversion of the geometry of a building starts from the stage of point cloud collection (Lidar technique) and later in their transformation through certain techniques and stages, into a voxel model. Classical plane segmentation algorithm is replaced by a voxel-based algorithm resulting from point clouds which proved to be more effective [17].

Lidar mapping: In order to achieve great detail and accuracy in indoor mapping of buildings, Lidar is the leader of all the techniques due to high density point clouds for the accurate representation of the geometry with the development of low-cost Lidar systems and sensors for collecting point clouds [18].

3D modelling techniques: Many relatively complex techniques of 3D modelling construction [8] have been implemented and used in comfort conditions volumetric visualization (Fig. 4). Indicative 3D modeling techniques [19], among others, are: (a) volume (Constructive Solid Geometry – CSG), (b) Occupancy Grid Mapping and voxels [20], (c) point clouds [21], (d) surfaces including polygons / NURBS splines [22], (d)

3D scanning, (e) Object's recognition [23], (f) 3D topology GIS features [24], (g) real-time remote 3D digital design and spatial simultaneous simulation [9, 25], (h) spatial interactive visualization and multimedia representation [3].



Fig. 4. Different methods of 3D representation of the same building [26]

Displays techniques are used in VR and AR systems depending on the purpose of the application. Some of the most well-known visualization technologies are CAVE, Head Mounted Displays (HMD), See-through Displays (HMD clones for seeing virtual image superimposed over the real world).

VRML and X-VRML are scene description languages describing a 3D world construction and providing convenient access to databases, object-orientation, parameterization, and imperative programming techniques [27].

3.2 Related Systems – Discussions

Many simulation systems use voxels optimization techniques for 3D volumetric reconstruction, such as "Voxel Hashing" allowing the projection of 2D or 3D objects into a 1D hash table [28], merging voxels by OpenCL processing data allows [29] or classifying 3D triangle meshes [30] (Fig. 5).

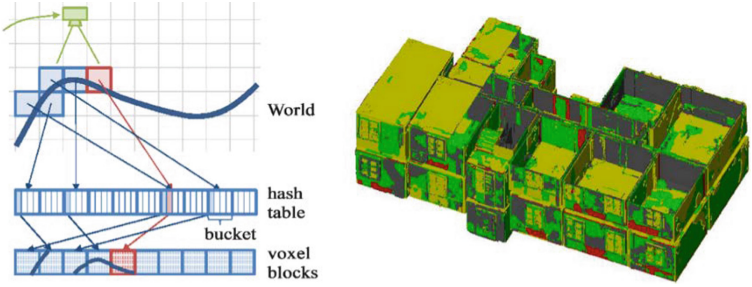


Fig. 5. Voxel-Hashing data structure. Source: [28], Voxel-based algorithm with 3D triangle meshes classification [30]

Other simulation systems incorporate AR into building management systems by using Tango Constructor app in order to capture spatial structures and texture information with a later processing of cloud dataset in MATLAB and a 2D temperature simulation in EnergyPlus software [5], (Fig. 6).

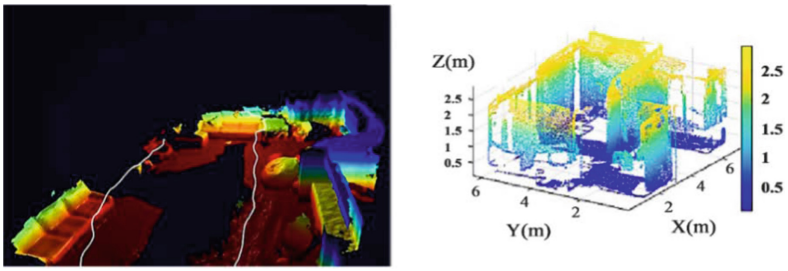


Fig. 6. Google's Project Tango technology maps the environment around it. Source: [31], Detected room 3D geometry. Source: [5]

Another system has been developed providing users with assistance for real-time monitoring of facilities' indoor thermal comfort conditions with VR [32] (Fig. 7) or AR technology with augmented virtual charts [2]. Other system integrates point cloud data and sensor data for visualization and analysis of a typical office building occupant comfort [6].



Fig. 7. Screenshot of the main interface of the prototype system with sensor outputs rates [32]

Finally, AR visualization systems of aerodynamics or fluid dynamic simulation use sensors and specific open-source software, such as Paraview in order to simulate an indoor air purification [33], or use sensors and library of speech and gesture recognition tasks in AR allowing 2D simulations [34] (Fig. 8).

3.3 Innovation of the Research Proposal

The proposed research is innovative as it focuses on a differentiated approach to representing and evaluating energy values from sensors inside buildings. It is a practical,

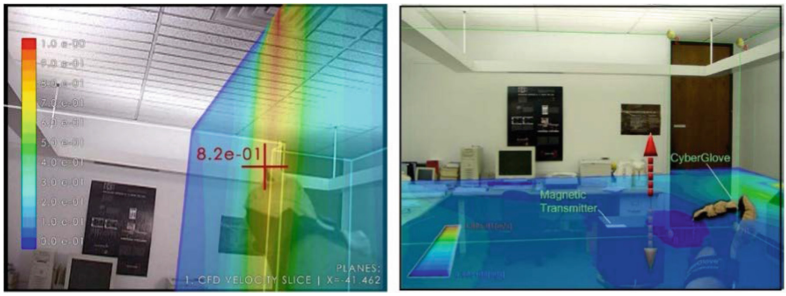


Fig. 8. Isoplane as seen by the user through the HMD and by the user using AR system [34]

pedagogical and immersive approach to understanding the interior structure of a building from an energy and health point of view. The innovation consists, among others, of an adapted method representing indoor energy values by sensors, adapted ways of building's modelling, values 3D representation, data processing and evaluation process, use and communication between new technologies and devices, and finally the social and scientific impact. In particular, the innovation of this proposal concerns:

(1) a simultaneous 3D building structure representation with an energy conditions values representation in 3D point clouds / voxels either in immersive VR of AR environment.
(2) a method to obtain accurate measurements by improving a grid installation of low-cost sensors system in adapted timing related to sensor rates evaluation for air temperature, humidity and pollution,
(3) an adapted methodology of volumetric building modelling based on BIM and energy information for MR comfort conditions representation
(4) provide future users / non-expert with the ability to real-time walk through and experience, through VR / AR environment the building's performance,
(5) a framework creation for making a Post Occupancy Evaluation (POE) of energy efficiency (EE) in buildings' indoor environment by comparing the measured results with optimal rates and offering an embedded interactive virtual 3D table-menu visualization in Virtual Environment during the indoor real-time walkthrough of a non-expert user,
(6) a significant contribution to a non-expert user for understanding "controversial comfort areas" and sensing appropriate actions based,
(7) a constantly re-feed base knowledge by the proposed representation system's use and experience of an expert or non-expert user,
(8) ensure an immersive pedagogical experience and awareness and of a non-expert user or an academic student on the appropriate comfort conditions through a volumetric interactivity virtual tour.

4 Scientific and Social Impact

4.1 Scientific Impact

The major scientific impact categorization of AR 3D representations concerns "Changes of proposed structures", "Design process", "Guiding visitors' experiences" and "Simulations of interdisciplinary information" [35].

The scientific impact of the proposed pilot system is significant because it allows the simultaneous representation of all elements with a real-time user's walkthrough inside the building's structure when the current simulation systems use 3D charts or 3D point clouds/voxel representations in a static virtual 3D building model inside a specific energy simulation software or 2D isoplane / 2D table simulation in a VR / AR building environment. The proposed system also creates special communication interfaces between point clouds/voxels techniques and Constructive Solid Geometry / BIM / GL Transmission / MR techniques and applications [35].

The proposed methodology provides very clear indications of the location, as well as the type of problems to be addressed by associating the classification building's interior geometry with its energy performance and air quality.

The contribution to the sustainability of energy efficiency in Buildings and the effective data manipulation with the octree structure of voxels for data mining and visualization of the measured values from the distributed sensors, play a crucial role since it can produce time-varying data.

Finally, the mobile low-cost system for extracting 3D building geometry and for measure environmental parameters such as (temperature, pressure, noise, air quality, etc) contributes to a cost-effective tool for innovative, more accurate and more understandable visualization, analysis and prediction of thermal comfort.

4.2 Social Impact

The proposed pilot system creates user awareness of the building's performance and air quality as until now assessing a building's EE and air quality was an activity for experts.

Association between internal geometric layouts and building energy performance and air quality will in turn provide designers with valuable knowledge with which they can fine tune their simulation tools.

Finally different partners target-groups can benefit from the proposed system [35], such as, Real estate market (high values to energy efficiency buildings), Engineers (simulations for energy demand of buildings), Customers (appropriate selection of apartment), Academic students (knowledge and evaluation), Public Authorities (contribution to a strategic plan for a social energy and health policy exertion / smart city platform's component for information, participation and deliberation).

5 Research Methodology & Pilot 3D Spatial Representation

5.1 Introduction

The research methodology of the present study is based on previous studies of the same authors on nosocomial infections detection [7] and Air Quality emissions from biosensors inside buildings [8].

The present research includes interdisciplinary approaches and techniques as they develop independently taking into account specific communication specifications regarding other techniques, the connection between them affects the overall development of the entire research organization. The methodological steps consists in (1) energy design through the selection of buildings where the measurements will be implemented based on geometry criteria, (2) comfort conditions measurements, implementation and temporary storage, (3) 3D building geometry's modelling in a MR environment, (4) 3D modelling of the selected comfort conditions values adapted to the building's geometry model by using special 3D point clouds / voxels techniques, (5) adapted evaluation of comfort conditions values, and (6) final 3D spatial representation of all previous elements in a unique MR interactive environment (Fig. 9).

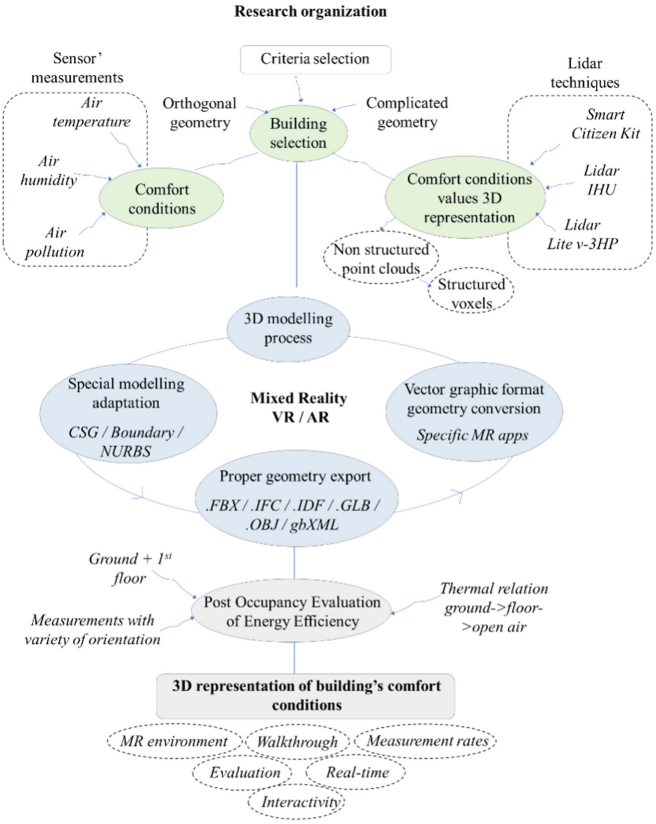


Fig. 9. Research schematic methodology

5.2 Buildings' Type Selection

Building use: the buildings to be measured are residences. They often represent the majority of the building sector that will very shortly need environmental certification in

order for users to proceed to any type of activity (renting, selling). It is also the type of building that has the least documentation on its envelope structure, energy efficiency and air quality, especially for those built before 1980. These buildings will require interventions in order to be certified that they comply with current energy efficiency regulations. Users will need to have awareness regarding both the energy efficiency as well as environmental factors that influence health issues.

Building geometry: in contrast with geometries from other building types (i.e. offices, schools) residences can display fairly complicated internal layouts. In order to examine the relation of internal geometry and indoor environmental factors, two buildings of varying interior geometry will be measured: one with simple orthogonal geometry and one with a complex configuration. This will establish our proposed method as suitable for any building types.

Relation with urban fabric: both buildings are free standing, consisting of a ground and first floor, thus providing measurements of spaces with a variety of orientations, as well as the thermal relation of ground-building, ground floor to first floor and first floor to open air.

Measurement conditions: the buildings will not be occupied, with systems of HVAC not in operation. Measurements will take place on a 24 h basis, in order to monitor value fluctuations. Measurements will take both place in winter (February) and summer (July), so as to test two opposite weather conditions. These will establish our proposed method for obtaining, visualizing and evaluating data for the actual building shell. Measurements can then take place after occupants move in and the building is in full use.

5.3 Sensors Implementation & Comfort Conditions Measurements

The current proposal will explore the above methods for extracting the geometry of buildings as well as collecting point clouds through the Lidar technique. The non-structured information provided by the point clouds will be transformed into structured voxels [36], which will be the framework for visualizing the various parameters such as temperature, humidity, air quality, noise, etc. that will be measured with low cost but high added value sensors. The sensors to be used in the research fall into two categories: (a) distance measuring sensors and (b) sensors for measuring environmental parameters. For the first category as mentioned, there will be used a low-cost lidar sensor (Fig. 10) suitable for indoor applications. For the efficient and effective use of such a sensor, a small platform will be constructed. For the collection of the geometric information (point clouds) of buildings, an effort will be made to develop and use a low-cost Lidar system [37], which will test the low-cost sensor LIDAR-Lite v3HP of Garmin company.

For measuring the environmental parameters, low-cost sensor such as the Smart Citizen Kit (smartcitizen.me), will be used, which has been developed as part of the iSCAPE project under the European Community's H2020 Program under Grant Agreement No. 689954 [38] (Fig. 10).



Fig. 10. Low-cost sensors: LIDAR sensor - Lite v3HP of Garmin company. Source: [32], and environmental sensor Smart Citizen Kit [33]

5.4 3D Building Geometry Modelling Process

Creating 3D spatial models for representation with 3D energy clouds points in an interactive MR environment is a quite complicated process. This is due, among others, to:

(a) the specific requirements of the MR representation regarding the communication protocols of the 3D objects,
(b) the plethora of 3D modeling and representation applications creating different formats and techniques,
(c) the involvement of energy values representation, special connected devices such as sensors, specific glasses, wireless technology, creating technical difficulties in their communication,
(d) the technology that governs the entire MR environment constantly evolving and creating new experimental communication standards.

The methods of creating 3D models vary in relation to the needs of each study and to the involvement and use of specific tools. The method of objects modelling photos is the process of creating a 3D object from scratch using dedicated specific 3D software. The photogrammetry method uses compilation of photos of the object from multiple angles (approximately 80–200) extracting 3D geometry and texture data by using 3D dedicated software. The Constructive Solid Geometry – CSG (CSG), Boundary and NURBS polylines method creates solid CAD models that must be converted by specific software into universally supported formats such as VRML [39], BIM file format, Collada, OBJect, etc. [40] within many cases specific model elaboration in order to optimize the converted models to comply with web/device standards [8]. To do so, these model formats are converted in proper format used across most platforms such as websites, Android and iOS. Indicative formats are GL Transmission Format (gLTF) or the accompanying gLB format which are royalty-free specification and common publishing open format for the efficient transmission and loading of 3D scenes and models by applications, including support for PBR (Physically Based Rendering) materials for realistic rendering and USDZ format which is Apple's optimized 3D file format for mobile devices using Universal Scene Description (USD) [41] format build in collaboration with Pixar. Finally, some MR applications (VR / AR) visualize 3D scenes after proper vector graphic formats transformation supported by specific 3D models creation software while specific

visualization equipment such as Tablet, smartphone, specific MR glasses (Hololens, Head Mounted Display - HMD), etc. [14] is needed for the final interactive 3D visualization. The selection of 3D modelling process and the corresponding techniques to be implemented in this work will be estimated and adapted to the research progress and the specifications needs of all element's integration such as sensors, clouds points 3D representation, etc.

5.5 Comfort Conditions Values 3D Representation

To generate building models directly from point cloud data (e.g. for engineering simulation) there are 3 dominant methods [42]: (a) Constructive Solid Geometry (CSG), (b) Boundary Representations (B-Reps) and (c) Spatial sub-division representations where an object domain is decomposed into cells with simple topological and geometry structure, such as regular grids and octrees.

The choice of one method than another depends on things like the nature of the application, the demanding processing power, the complexity of the object and others. The current study deals with the third dominant method, which is based on voxel segmentation. Voxels can represent complex objects in reducible, discrete units (like particles), they can be incredibly powerful for simulating real world behavior of complex objects. The main advantages of voxels are:

(a) They are a more "accurate" 3D building block than any other modeling type, as they mimic particles,

(b) They unlock new simulation techniques that would be impossible with other modeling methods,

(c) Voxels are the quickest way to quickly model and visualize volumetric data (especially in natural or organic formations).

The 3D representation procedure to be followed is: (1) Lidar scans of a room or building, (2) Registration of different scenes, (3) Cleaning process of unwanted points, (4) Rendering, (5) Export of unstructured point clouds (.las files), (6) Filtering of point clouds through voxel-based sampling filters (e.g. Voxel Center Nearest Neighbor filter), (7) Import of structure point cloud (voxel) in open source visualization software, (8) Visualization of the model with different voxel methods (slice etc), (9) Integrate the sensors values (temperature, humidity, etc) into the voxel model and visualize the interpolation values and (10) Export of the visualized model for use in AR, VR etc. applications.

5.6 Comfort Conditions Values Evaluation

Our proposal seeks to establish a method for visualizing indoor environmental data as well as air quality. The methodology will initially be established for primary data such as air temperature, relative humidity, noise pollution, PM (air pollution), as they are environmental factors that are of primary importance to users' health. This methodology can further be expanded to include more factors, such as MRT, thermal comfort, EMR radiation. After measurements are taken and 3D representations are constructed, our proposal presents users with an evaluation mode. In order to choose the best type of

evaluation visualization, a pilot study with questionnaires for users can be performed, investigating their preferred evaluation type. A radar chart (if chosen) will be assigned to each of the parameters measured, providing instantly to the user, while walking through the space, with the value requested, located in the chart in relation to the ideal comfort zone.

Air temperature is the primary environmental factor that users can easily perceive. Also, the percentage of relative humidity on a daily basis, throughout the year is a decisive factor in indoor air quality, as high percentages indicate environments potentially dangerous for occupants' health, facilitating the existence of airborne allergens and bacteria, among others.

Noise pollution is another crucial factor contributing to occupants' health. Anxiety, poor sleep and performance and hearing impairment have been registered as possible negative effects of continuous exposure to high levels of noise in the building envelope [43]. Measurements will be taken with windows and doors closed. A clear picture of the building's performance is formed with measurements on a daily and weekly basis, while the building is not occupied and while in full operation.

Air quality is monitored, among others, through the measurement of particle matter (PM). In particular particles with an aerodynamic diameter of 2.5 μm or smaller (PM2.5) pose a significant threat to human health, as they can penetrate deeply into the lungs and cause serious damage that triggers potentially lethal diseases [44]. The WHO have set standards for PM2.5 concentration: 25 μg m–3 (24-h mean) and 10 μg m–3 (annual average) [45]. The air pollution index for PM2.5 is divided into six levels: 0–50, 51–100, 101–150, 151–200, 201–300, and greater than 300. Based on these the severity of air pollution in an environment can be specified [46]. Indoor PM concentration is a combination of factors related to outdoor environment as well as occupants' activities indoors. PM2.5 concentration visualization can provide occupants with a clear understanding of the surrounding outdoor air quality, as well as the effects of their own actions, when measurements are taken when the building is in use. Comparison of measurements from buildings that have a different internal geometric layout can further reveal if a complex geometry can work in favor or against PM concentration in a residence.

5.7 Pilot 3D Spatial Representation of Building's Comfort Conditions in MR Environment

Technical approach: The modeling as well as the representation of the building's geometry types will be implemented either in VR and / or AR environment. The technical approach of 3D modeling as well as the Level of Details (LOD) will be adapted in relation to the integration of the other elements, the communication with them, the technical limitations and the detection of the respective ideal way of elements and information revelation. The 3D building's geometry may be accompanied by a different type and level of customized rendering in relation to the representation environment, the memory and speed of the output devices equipment and the representation of the selected elements. The representation of the application in VR environment will need special VR head glasses and corresponding hand sensors, while the representation in AR environment will need a special smartphone / tablet with integrated tracking movements sensors. In

VR environment, 3D building's modelling is necessary as it is the "shell" (spatial reference) in which comfort conditions values and their evaluation will be simulated. In AR environment, building's 3D modelling is an additional but not necessary information for the 3D representation of comfort conditions values since the user can navigate the real space of the building (Real Reality) by seeing the space through a special AR device (smartphone / tablet) and at the same time simulate through the same device these values and their evaluation inside the existing building (Augmented Reality). The 3D representation of these, almost transparent due to their sparse density, values will be in the form of 3D points clouds / voxels that will allow their visual adjustment in space and the corresponding spatial perception of the user.

By using a point-based distance sensor (such as the LiDAR Lite v3HP) together with two servos for both the azimuth and elevation axes, a point cloud data is collected and can be exported to a comma separated value (CSV) file and geometry definition file (OBJ) [47]. This system will be used for the extraction of the 3D geometry (point clouds) of the room/building. Regarding the sensors for measuring the environmental parameters, the smart citizen kit is designed with the ability to expand, with a central data recorder and the ability to connect to the Internet. All of the above is based on open-source hardware and code.

Visualization results: The indoor pilot 3D representation of comfort conditions in MR environment will be implemented in VR and / or AR environment. In VR environment, 3D building's modelling process is necessary to integrate comfort conditions clouds 3D representation with the corresponding rates simulation and evaluation. Special VR head glasses and hand sensors will be needed for real-time walkthrough. In AR environment, 3D building's modelling is not necessary and it be replaced by the building's Real Reality (RR). Special AR devices (smartphone / tablet) with integrated trackers / sensors simulate augmented comfort conditions clouds in superimposed transparent representation with augmented objects of rates charts and evaluation. In both cases, 3D representation of spatial and environmental phenomena presents a high degree of spatial perception (Fig. 11).



Fig. 11. Visualization results of the proposed pilot representation system in MR environment

6 Conclusions

This research attempts a 3D spatial representation and ergonomic assessment of energy and health design inside a building. The final specific interactive 3D architectural representation of energy and health conditions rates, in relation to the building's 3D spatial

simulation in a virtual environment of mixed Reality (Real + Virtual + Augmented Reality), occasionally adapts different representation techniques so that the final representation becomes perceived by an average habitant, through an architectural, spatial representative, ergonomic and communicative approach.

The research methodology involves, among other, buildings' energy design based on geometry criteria, comfort conditions measurements, 3D building geometries and comfort conditions modelling and adaptation in a MR environment and final 3D spatial representation of all previous elements in a unique MR interactive environment.

The scientific and the social impacts are very significands because on one hand all sensors can be combined via the internet and with in situ measurements, to transfer the results of the measurements to a virtual environment and on the other hand the contribution to a strategic plan for a social energy and health policy exertion by public authorities via implementation of the proposed visualization / evaluation system to selected buildings' local network.

The integrated technical approach of the proposed system allows the use of special VR head glasses and the corresponding hand and point-based distance sensors and the comfort conditions values in the form of 3D points clouds / voxels. The visualization results allow the integration of comfort conditions clouds 3D representation with the corresponding simulation and evaluation rates and the use of special AR devices (smartphone / tablet) with integrated trackers / sensors that simulate augmented comfort conditions clouds in superimposed transparent representation with augmented objects of rates charts and evaluation.

Ensuring an immersive pedagogical experience, awareness and information of a non-expert user or an academic student on the appropriate comfort conditions through a volumetric interactivity virtual real-time walkthrough, the perspectives of the proposed system are very promising. This is due to, inter alia, the dissemination of methodology and evaluation methods, through assessment of users' reception of the visualization and evaluation of indoor environmental data with surveys and questionnaires and finally to an adjustment of the proposed system to an accessible online platform as a component of a smart city platform for public awareness, information, participation and deliberation.

References

1. Chang, K.M., Dzeng, R.J., Wu, Y.J.: An automated IoT visualization BIM platform for decision support in facilities management. Appl. Sci. **8**, 1086 (2018)
2. Natephraa, W., Motamedib, A.: Live data visualization of IoT sensors using augmented reality (AR) and BIM. In: Proceedings of the 36th International Symposium on Automation and Robotics in Construction-ISARC (2019)
3. Kouzeleas, S.: Processes aspects of modeling, interactive digital visualization and multimedia representation of architectural and urban area. In: Proceedings of the 11th International Conference on Computer Graphics and Artificial Intelligence (3IA'2008) May 30–31, 2008, Athens, Greece, pp. 205–210 (2008)
4. Mahdavi, A., Brahme, R., Gupta, S.: Self-aware buildings: a simulation-based approach. In: Proceedings of the Seventh International IBPSA Conference, Rio de Janeiro, Brazil (2001)
5. Aftab, M., Chau, S.C., Khonji, M.: Enabling self-aware smart buildings by augmented reality. In: Proceedings of the Ninth International Conference on Future Energy Systems, pp. 261–265 (2018)

6. Stojanovic, V., Trapp, M., Hagedorn, B., Klimke, J., Richter, R., Dollner, J.: Sensor data visualization for indoor point clouds. In: Proceedings of the 15th International Conference on Location Based Services (LBS 2019), Vienna, Austria (2019)
7. Kouzeleas, S., Tsolakidis, I.: Contribution to nosocomial infections detection via biosensors particles spatial simulation in mixed reality environment. In: Computational Science and its Applications – ICCSA 2025 Workshops. ICCSA 2025 Lecture Notes in Computer Science, vol. 15891, pp. 17–32 (2025)
8. Kouzeleas, S., Tsolakidis, I.: Assessment contribution of an architectural indoor healthy status via biosensors particles spatial simulation. In: Computational Science and its Applications – ICCSA 2024 Workshops. ICCSA 2024 Lecture Notes in Computer Science, vol. 14825, pp. 168–183 (2024)
9. Kouzeleas, S.: Architectural and urban spatial digital simulations. Int. J. Eng. Model. **36**(2), 95–118 (2023)
10. Li, W., Nee, A., Ong, S.K.: A: state-of-the-art review of augmented reality in engineering analysis and simulation. Multimodal Technol. Interact. **1**(3), 17 (2017)
11. Milgram, P., Kishino, F.: A taxonomy of mixed reality visual displays. Trans. Info. Syst. **12**(12), 1321–1329 (1994)
12. Kouzeleas, S., Nikolaidou, S., Goussios, D., Goulas, A.: Pilot interactive visualization tool of a participatory guarantee system: the case of Terra Thessalias PGS. Int. J. Eng. Innov. Technol. **9**(9), 1–17 (2020)
13. Baratoff, G., Blanksteen, S.: Tracking devices. In: The Encyclopedia of Virtual Environments (1993)
14. Zlatanova, S.: Augmented Reality Technology. GISt Report 17. Delft (2002)
15. Goldstein, R., Breslav, S., Khan, A.: Towards voxel-based algorithms for building performance simulation. In: Proceedings of the IBPSA-Canada eSim Conference (2014)
16. Xin, C., Hui, Z., Yueqi, H., Zhaoyang, W., Jinong, D., Kaichao, W.: Real-time 3D visualization of temperature distribution for data Center thermal management. In: Proceedings of the 3rd International Conference on Multimedia Technology (ICMT-13), pp. 895–902 (2013)
17. Huang, M., Wei, P., Liu, X.: An efficient encoding voxel-based segmentation (EVBS) algorithm based on fast adjacent voxel search for point cloud plane segmentation. Remote Sens. **11**(23) (2019)
18. Thomson, C., Apostolopoulos, G., Backes, D., Boehm, J.: Mobile laser scanning for indoor modelling. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. **2**(5W2), 289–293 (2013)
19. Schlaug, F.: 3D Modeling in Augmented Reality Thesis. Department of Electrical Engineering, Linkoping University, Sweden (2013)
20. Elfes, A.: Using occupancy grids for mobile robot perception and navigation. Computer. **22**, 46–57 (1989)
21. Qu, T., Sun, W.: Usage of 3d point cloud data in bim (building information modelling): current applications and challenges. J. Civ. Eng. Archit. **9**(11), 1269–1278 (2015)
22. The Apple Technical Journal: NURB Curves: A Guide for the Uninitiated, Homepage (2025). https://download.blender.org/source/chest/blender_2.03_tree/docs/nurbs.html. Accessed 30 Oct 2025
23. Russell, S.J.: Norvig: Artificial Intelligence: a Modern Approach, Book, 2nd edn. Prentice Hall (2003)
24. Zlatanova, S.: 3D modelling for augmented reality. Int. Arch. Photogramm. Remote Sens. **34**(2W2) (2001)
25. Kouzeleas, S.: Real-time remote monitoring urban and architectural 3D digital design and spatial simultaneous simulation. Int. J. Adv. Res. **6**(12), 360–377 (2018)

26. Poux, F.: How to represent 3D data, Homepage (2020). https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://orbi.uliege.be/bitstream/2268/248375/1/How%2520to%2520represent%25203D%2520data.pdf&ved=2ahUKEwjO9sWI8pqUAxV6nf0HHX6vMVoQFnoECB8QAQ&usg=AOvVaw0RfGwNG5vwk5-HvVRJVb7v. Accessed 5 Nov 2025
27. Walczak, K., Cellary, W.: Building database applications of virtual reality with X-VRML. In: Proceedings of the 7th International Conference on 3D Web Technology, Web3D'02, USA (2002)
28. Nie, M., Zollhofer, M., Izadi, S., Stamminger, M.: Realtime 3d reconstruction at scale using voxel hashing. ACM Trans. Graph. **32**(169), 1–11 (2013)
29. Hui, Z., Xu, Y., Wu, Y.: Multi-scale voxel hashing and efficient 3d representation for mobile augmented reality. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops (2018)
30. Hübner, P., Weinmann, M., Wursthorn, S.: Voxel-base indoor reconstruction from hololens triangle meshes. In: Proceedings of the Annals of the Photogrammetry Remote Sensing and Spatial Information Sciences, V-4-2020, pp. 79–86 (2020)
31. Williams, R.: What is Google's project Tango? Homepage (2015). https://www.telegraph.co.uk/technology/google/11391989/What-is-Googles-Project-Tango.html
32. Natephraa, W., Motamedib, A.: Bim-based live sensor data visualization using virtual reality for monitoring indoor conditions, intelligent & informed. In: Proceedings of the 24h CAADRIA Conference, Victoria University of Wellington, New Zealand, pp. 191–200 (2019)
33. Kim, M., Yi, S., Jung, D., Park, S., Seo, D.: Augmented-reality visualization of aerodynamics simulation in sustainable cloud computing. Sustainability. **10**(5), 1362 (2018)
34. Malkawi, A.M., Srinivasan, R.S.: A new paradigm for human-building interaction: the use of CFD and augmented reality. Autom. Constr. **14**(1), 71–84 (2004)
35. Kouzeleas, S.: Scientific impact of architectural indoor augmented reality 3D digital representations. Int. J. Adv. Res. **10**(02), 425–432 (2022)
36. Gebhardt, S., Payzer, E.: Polygons, point-clouds, and voxels, a comparison of high-fidelity terrain representations. In: Proceedings of the Fall Simulation Interoperability Workshop, pp. 357–365 (2009)
37. Athavale R., Harish Ram D.S., Nair B.B.: Low-cost solution for 3D mapping of environment using 1D LIDAR for autonomous navigation. In: Proceedings of the IOP Conference Series: Materials Science and Engineering 561(1) (2019)
38. Alvarez-Campana, M., López, G., Vázquez, E., Villagrá, V.A., Berrocal, J.: Smart CEI moncloa: an iot-based platform for people flow and environmental monitoring on a smart university campus. Sensors. **17**(4) (2017)
39. Lorenz, M., Spranger, M., Riedel, T., Pürzel, F., Wittstock, V., Klimant, P.: CAD to VR – a methodology for the automated conversion of kinematic CAD models to virtual reality. In: Proceedings of the Procedia CIRP, vol. 41, pp. 358–363 (2016)
40. Barazzetti L., Banfi F.: Historic BIM for Mobile VR/AR applications Ioannides, M., Magnenat-Thalmann, N., Papagiannakis, G. , in: Mixed Reality and Gamification for Cultural Heritage, pp. 271--290 (2017)
41. Pixar Animation Studios: Introduction to USD (Universal Scene Description), homepage (2021). https://openusd.org/release/intro.html
42. Hinks, T., Carr, H., Truong-Hong, L., Laefer, D.F.: Point cloud data conversion into solid models via point-based voxelization. J. Surv. Eng. **139**(2), 72–83 (2013)
43. Pujol, S., Berthillier, M., Defrance, J., Lardies, J.: Indoor noise exposure at home: a field study in the family of urban schoolchildren. Indoor Air. **24**(5), 511–520 (2014)
44. Hegde, S. et al.: Indoor household particulate matter measurements using a network of low-cost sensors. Aerosol Air Qual. Res. **20**, 381–394 (2020)

45. World Health Organization: WHO global air quality guidelines, homepage (2021). https://www.who.int/publications/i/item/9789240034228
46. Sangik, L., SeoKyung, P., Jaeyoung, S., Hyunsoo, L., Jin-Kook, L., Hyunjung, K.: Visualization of the airborne dust concentration on the building floor plans. In: Proceedings of the 32nd International Symposium on Automation and Robotics in Construction and Minings – ISARC, Finland, pp. 516–521 (2015)
47. Shahrin, M.R., Hashim, F.H., Zaki, W.M.D.W., Hussain, A., Raj, T.: 3D indoor mapping system using 2D LiDAR sensor for drones. Int. J. Eng. Technol. **7**(4), 179–183 (2018)

Smart City Digital Twin Framework for Semantic Indexing and Dynamic Modelling of 3D Buildings

Marco Fanfani[1(✉)], Federico Galli[1], Chiara Garau[2], Paolo Nesi[1], and Gianni Pantaleo[1]

[1] DISIT Lab, DINFO, University of Florence, Florence, Italy
{marco.fanfani,federico.galli,paolo.nesi,
gianni.pantaleo}@unifi.it
[2] DICAAR, University of Cagliari, Cagliari, Italy
cgarau@unica.it

Abstract. Smart City Digital Twins (SCDT), including 3D representations of the urban environment, have gained ever-increasing attention from academic, public, and private sectors in the last years. Being faithful, dynamic, three-dimensional replicas of the cities, SCDTs realize concrete decision support systems to facilitate the work of urban planners and decision makers. Over the years, several municipalities have succeeded in producing 3D city maps, particularly useful in managing and planning the development of the built infrastructure. However, in most cases such representations are static replicas without any semantic enrichment and offer very limited interactive visualization and manipulation, thus inadequate to support what-if analysis. In this paper, the 3D city modelling and visualization framework embedded into the Snap4City Digital Twin platform is proposed. Each building is represented with multiple 3D versions, augmented with semantic information, and modelled as an individual entity by exploiting the KM4City ontology. Web-based 3D dashboards are used to present the urban twin and allow users to inspect, control and modify in real-time the 3D virtual environment in order to evaluate different scenarios for development of the urban spaces.

Keywords: 3D Building · Knowledge Base · Ontology · Smart City · Digital Twin

1 Introduction

In recent years, Smart City Digital Twins (SCDTs) have evolved from static 3D city visualization tools to dynamic, data-driven decision support systems that integrate heterogeneous urban data, real-time streams, and interactive analytics into a unified digital environment [1–3]. Within SCDTs, buildings are central because they connect to mobility, energy, environment, and service organization, serving as primary units for scenario analysis and policy testing and representing the physical backbone through which many urban processes become observable, measurable, and manageable [4, 5].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 21–34, 2027.
https://doi.org/10.1007/978-3-032-30539-8_2

However, many existing 3D city models remain largely visual, with limited semantic enrichment and interactive capabilities, and are therefore still used mainly as descriptive representations rather than as operational components of urban intelligence systems, hindering scalable decision support at the building level. At the same time, the growing availability of geospatial data, IoT information, and multi-source urban datasets makes it increasingly necessary to move beyond purely geometric modelling approaches and toward solutions capable of linking 3D objects with structured knowledge and dynamic attributes. These gaps motivate the need for a framework capable of combining semantic indexing, dynamic 3D modelling, and interactive visualization in a coherent and scalable way, so that buildings are no longer treated as static graphic objects, but as semantically enriched entities that can be queried, updated, compared across temporal states, and manipulated to explore alternative urban scenarios. In this context, the objective of this article is to investigate the following key research question: how can 3D building representations within SCDTs be semantically enriched and dynamically managed in order to support scalable, interactive, and data-driven urban decision-making? To address this question, the paper proposes a framework embedded in the Snap4City Digital Twin platform, where buildings are modelled as semantic entities through the KM4City ontology, associated with multiple 3D representations, and made accessible through interactive web-based dashboards for real-time inspection and scenario analysis. More specifically, the proposed approach aims to bridge the gap between semantic interoperability, large-scale building data management, and user-oriented 3D interaction, thereby supporting not only the visualization of urban space but also the exploration of possible transformations and what-if scenarios. To achieve this, the paper is organized as follows: Section 2 presents the semantic building modelling approach and the entity creation process within the Snap4City platform. Section 3 describes the dynamic and interactive 3D visualization framework and its main functionalities, while in Sect. 4 validation activities are reported to demonstrate the concrete applicability of the proposed framework. Finally, Sect. 5 draws the conclusions and discusses possible future developments.

2 Semantic Building Modelling and Entity Creation

In this section, the Snap4City smart city digital twin framework (https://snap4city.org) [2] is presented, focusing on building representation and management. First, the Snap4City platform is briefly introduced, illustrating its main functionalities. Then details on the solution used to model building entities into Snap4City are discussed, presenting their semantic indexing according to the KM4City ontology [6, 7] and the defined data model. Finally, the ingestion process implemented with processing logic embedded into the platform is described.

2.1 Snap4City Platform

Snap4City is an open-source platform designed to realize and implement Smart Cities Digital Twin (SCDT) in order to provide a decision support system helping decision makers in operational monitoring activities and in strategic and tactical planning of the urban environment. The platform is able to integrate and manage heterogeneous data

from multiple sources guaranteeing maximal interoperability, and, being based on a microservice architecture, it offers wide scalability to allow its deployment in different sized cities, from small municipalities to big metropolitan areas. Ingested data is modelled into semantic entities based on classes defined in the KM4City ontology. Semantical indexing is sided with multiple big data storages (i.e., NoSQL databases) in order to store the ingested information. Specific APIs have been developed to allow fast data retrieval via relational, semantic and geographical queries. Processing logic and analytical services are available to process data and perform complex computations exploiting simulations engines and artificial intelligence solutions. Interactive dashboards are included to present data to the users and allow them to inspect and study the urban environment.

The architecture of Snap4City is based on multiple functional layers that cooperate to provide a complete solution for urban data management and analysis. The platform aims to offer high-quality services across several development phases.

Acquisition, Indexing and Storage: Manages the semantic ingestion of heterogeneous data from IoT sensors, municipal databases, open data, web services, and other sources [8], ensuring interoperability through the adoption of established standards, such as FIWARE Smart Data Models and IoT communication protocols.
Processing: Implements advanced analytical capabilities to produce high-level information including, for example, reconstruction of traffic density information from scattered data [9], prediction of air quality [10], assessment of public and sharing transportation performance [11], thus enabling the possibility to carry out what-if analyses to evaluate the impact of changes within the urban environment.
Visualization and Distribution: Creation of interactive 2D and 3D web interfaces accessible via standard browsers [12], which allow local authorities and citizens to explore the city's Digital Twin visualizing real-time and historic data and simulate alternative scenarios with high precision [13].

2.2 Modelling Building into the Snap4City Platform

The Snap4City platform is built upon the KM4City ontology. Smart city ontologies [14] play a pivotal role in both allowing semantic enrichment and supporting data interoperability [15]. Differently from other state-of-the-art ontologies for the smart city domain [16–20], in which entities usually are modelled with an articulated and overcomplex hierarchy of classes leading to an explosion of instances when a huge amount of data must be ingested and limiting the possibility to accommodate different kinds of data without frequent reengineering, KM4City adopted a different approach to facilitate the deployment and exploitation of semantic representation in SCDT platform.

Focusing on building representation, the *km4c:Building* class has been derived from a general *km4c:HighLevelType* (HLT) class. HLT is a broad concept that can be used to describe different kinds of entities relevant in the smart city context. It can be seen as a sort of label to add semantic information to data of different kinds. In addition, to support interoperability, the *km4c:Building* has also been subclassed by *s4city:CityObject* class from the SAREF4CITY ontology [21] (extension for smart cities of the IoT ontology SAREF [22]) and made equivalent to *bot:Building* of the Building Topology Ontology [23]. In Fig. 1**,** a simplified representation of the KM4City ontology is provided.

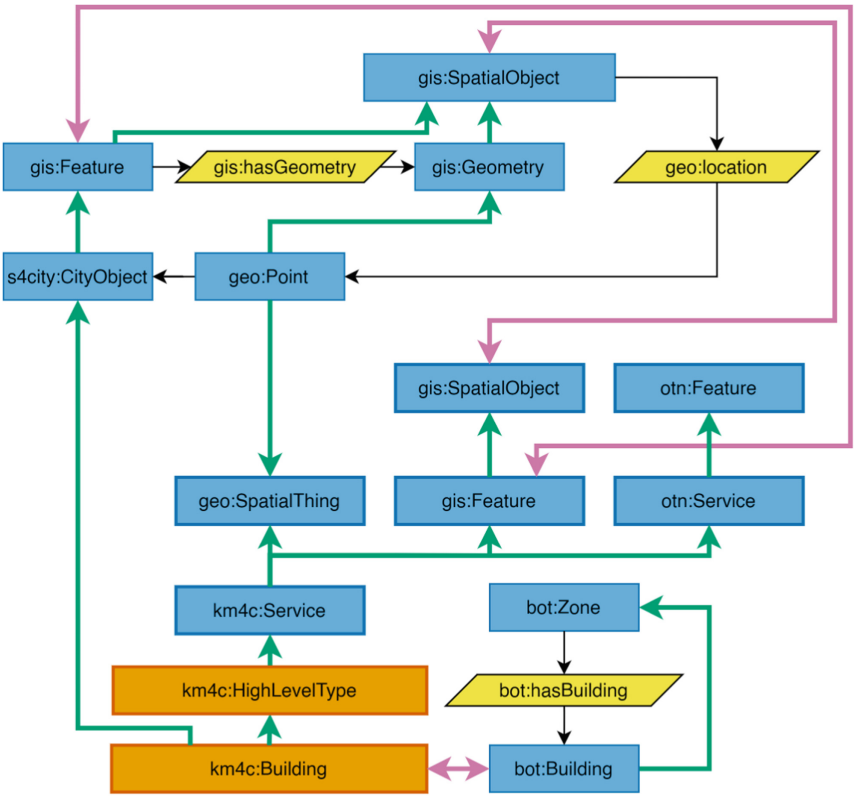


Fig. 1. Simplified diagram of the KM4City ontology focusing on *HighLevelType* and *Building* classes. Rectangles indicate classes, parallelograms are properties. Black arrows link triple relations, i.e., subject-property-object. Green arrows indicate superclass/subclass relations, while purple arrows show class equivalence. Prefixes are km4c for KM4City, bot for Building Topology Ontology, geo for WGS84, gis for GeoSPARQL, otn for Open Transport Network, s4city for SAREF4CITY.

Building entities are enriched with a set of additional data according to a defined data model composed by a set of static attributes and a number of variables each one semantically described with specific value type, value unit, and data types. For example, the address variable has value type *Geolocation*, value unit *text*, and data type *string*. In Table 1**,** the complete list of variables included in the data model of buildings, is reported.

Such data model allows to provide useful information for each specific building of the city, including its location (e.g., *address*, *city*, *postCode*), classification (*name*, *description*, *buildingType*), structural information (*numberOfFlats*, *numberOfLevel*, *numberOfUndergroundLevel*, *capacity*, *height*, *groundShape*, *isAccessible*), external sources (*osmID* from OpenStreetMap, *website*, *wikidata*, *wikipedia*), and available 3D models (*3DModel*, *3DModelType*, *3DModelDescription*). Note that multiple 3D models can be

Table 1. Variables of the building data model

Variable Name	Value Type	Value Unit	Data Type
dateObserved	Timestamp	ms	string
city	Geolocation	text	string
address	Geolocation	text	string
postCode	Identifier	ID	string
name	Identifier	ID	string
description	Status	text	string
buildingType	Status	coded status	string
numberOfFlats	Count	number	integer
numberOfLevel	Count	number	integer
numberOfUndergroundLevel	Count	number	integer
capacity	Count	number	integer
height	Status	number	float
groundShape	Geometry	complex data	JSON(GeoJSON)
isAccessible	Status	coded status	string
osmID	Identifier	ID	string
website	URL Link	webpage	string
wikidata	URI Link	ServiceURI	string
wikipedia	URL Link	webpage	string
3DModel	Identifier	ID	string
3DModelType	Status	coded status	string
3DModelDescription	Status	text	string

associated with the same building entity: for example, it is possible to have multiple representations at different Level of Detail (LoD), extending the classification of CityGML [24] by including also Building Information Models (BIM). In addition, associating multiple 3D representations enables the system to show temporal evolution of the same buildings, e.g., to monitor the progress of a construction site, or possible future changes yet to be built in the real environment, thus supporting urban development activities. Indeed, according to the data model, each building entity must include a *dateObserved* variable indicating the temporal instant in which the model refers to. In this way, it is possible to ingest the same building multiple times accumulating its evolution and possible future plans.

2.3 Massive Building Data Ingestion Pipeline

The main challenges in creating a urban digital twin capable to represent each single building as an independent entity are: (i) the data volume (a medium sized city like Florence, Italy, counts more than 50,000 buildings) requiring adequate platform architecture as well as fast and fault-tolerant ingestion pipelines able to recover from possible errors without requiring a full restart, (ii) the necessity to merge and reconciliate multiple data sources with robust mechanism to deal with missing data, and (iii) a visualization engine able to support massive rendering of 3D models. While characteristics of the rendering solution implemented into the Snap4City dashboards are described in Sect. 3, hereafter details on the ingestion pipeline are reported.

In the presented solution, two main sources of data are considered: OpenStreetMap (OSM), to recover building information and metadata, and a set of 3D building models reconstructed from Digital Terrains and Surface Models (DTM/DSM) following the method described in [25]. In particular, a local copy of OSM data have been used and instantiated into a Postgres database to allow faster data retrieval. Differently, 3D building models have been organized in a hierarchical structure mirroring the tiling system typically used in web maps. To allow data reconciliation each building model has been named with a prefix linked to the relative OSM ID.

Data ingestion has been implemented exploiting Snap4City Processing Logic/IoTApp (PLg) [26]. Snap4City PLg extends the Node-RED framework [27] with a series of specific functionalities (distributed as Node-RED libraries microservices) allowing full integration with platform services via APIs. Two different ingestion flows have been developed according to the complexity of the building: simple building, having a plant representable with a single polygon and indicated as *way* in OSM, and complex building, identified by the OSM *relations* and usually having plants composed by multiple polygons. Such a distinction is necessary to properly handle the different modalities in which OSM data are organized in the Postgres database for *ways* and *relations* and to manage the construction of the building plant geometry, eventually stored in the *groundShape* model variable (see Table 1) and into a knowledge base entity attribute in WKT format. As an example, the PLg flow for ingesting *way*-type buildings is shown in Fig. 2 and hereafter discussed.

The processing flow is composed of a sequence of steps (highlighted with gray boxes in Fig. 2). First, in step #0, variables to track the flow progress are initialized and available 3D models not having a match in the OSM database are identified and isolated. Eventually, such models will be used to create entities by providing required metadata manually.

After this first filter, in step #1, a batch of building IDs is selected and sent to the Queue Manager node. Such node keeps a queue of buildings to be processed and is responsible for releasing the next ID only after the ingestion process of the previous one is concluded (successfully or unsuccessfully). For each building ID its geographic coordinates are extracted from the OSM database in order to define its position and create its WKT geometry (used as shape representation in the knowledge base so as to enable geographic queries). Such data is used to define the building entity and, in step #2, is sent to the IoT Directory of the Snap4City platform to register the entity into the knowledge base. Once this operation successfully returns, in step #3 two operation are carried out.

First, actual building data are retrieved from the OSM database and inserted in the entity previously created, following the data model presented in Table 1**,** by sending the data to the internal context broker, thus initiating the ingestion process to store the information into a NoSQL database [8]. In parallel, the access policy of the entity is changed. Being a secure and privacy compliant platform, Snap4City enforces user ownership on any ingested data. By default, created entities are private and accessible only by the creator. Nevertheless, it is possible to make the entity public or delegate it in read, write, or read-and-write to other users. In the context of this work, all building entities have been made public.

To guarantee system robustness, the ingestion flow includes dedicated error-handling mechanisms with specific logging to register errors happening during the process. The managed error scenarios are three. The case of buildings without associated 3D models, shown in block E1 of Fig. 2; the failure of the IoT Directory during entity registration (illustrated in E2), and errors during data transmission to the context broker and subsequent data ingestion, highlighted in block E3. In the first case (E1) the flow logs the error and skips the specific building ingestion, since the 3D model is missing, it would not be possible to show it in the in 3D dashboard. In the other two cases recovery strategies have been devised. After a first failure, the operations are executed again for a given number of times after a short delay (set to 20 s) by re-running step #2 in case of error during entity creation, or step #3 in case of failed data ingestion. This is motivated by the fact that sometimes errors are related to temporary network failures, thus a simple retry could solve the problem. When such a strategy is not sufficient and the error persists, the failure is logged in order to be recovered in a successive time. Note that, in these cases, all the flow is halted for 10 min before moving on the next building. This solution has been designed to prevent the system from continuing sending requests in case of service problems in the IoT Directory to avoid overloading the platform infrastructure further.

Once all the batches have been executed, two different recovery strategies are executed to solve the logged errors. The first strategy, shown in block R1, is designed for recovering from errors that happened during entity creation: in this case steps #1, #2, and #3 all re-executed. Differently, errors during data ingestions are handled with block R2 that limits the re-execution to step #3. The recovery strategies start automatically for the first time. If failures happen again, these are logged by overwriting the error logs and the flow is stopped (exploiting a flow variable). This solution has been adopted since such repeated errors are probably due to some problems with the input data. Therefore, it is required that the operator correct the input and then manually restart the recovery flow and complete the ingestion. Note that both recovery strategies work on batches of IDs: while R1 injects the flow messages into step #1 exploiting its Queue Manager, R2 requires a different Queue Manager node, since steps #1 and #2 must not be executed in this case.

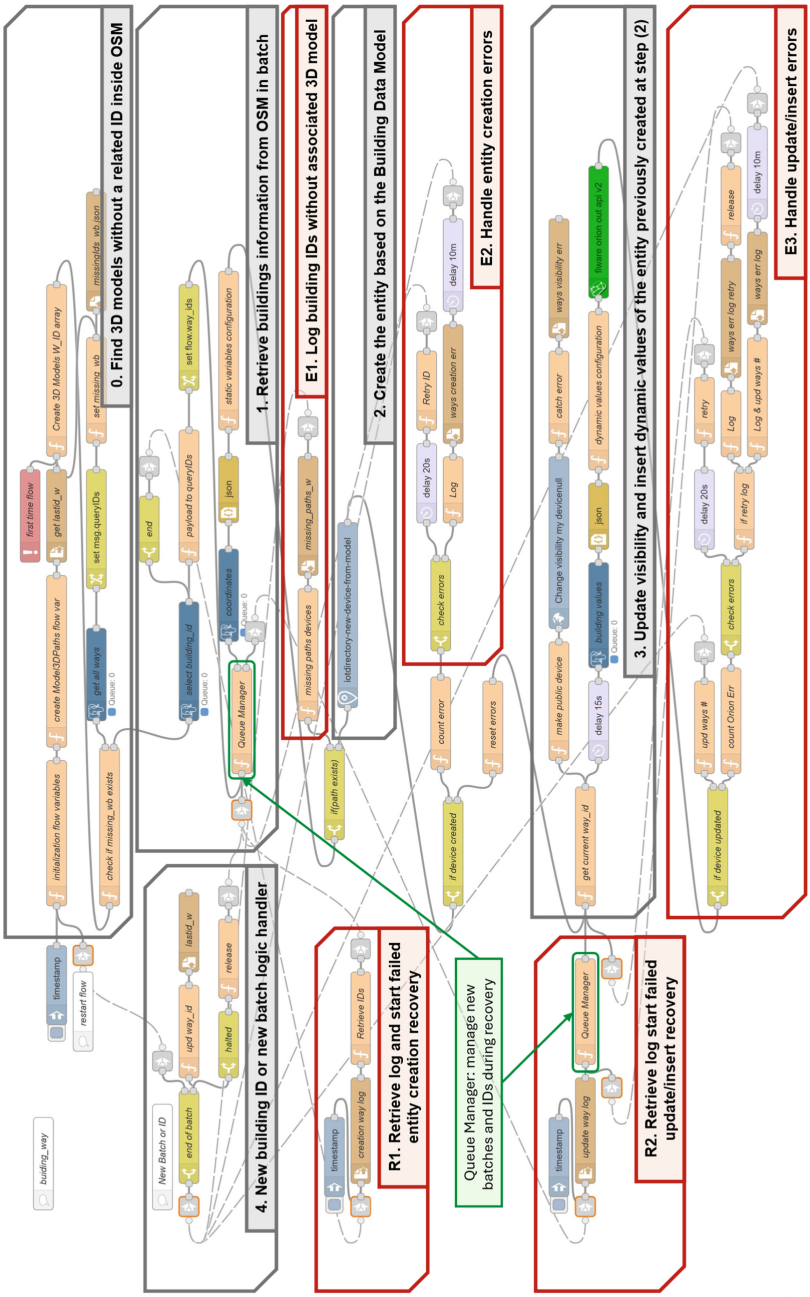


Fig. 2. Processing Logic ingestion flow for OSM ways. Different steps of the pipeline are highlighted with gray boxes and short descriptions. Differently sub-flows for error handling and recovery are highlighted in red.

3 Dynamic and Interactive 3D Visualization

The 3D representation of the urban Digital Twin is implemented through interactive dashboards in the Snap4City platform. In particular, the proposed solution exploits the Snap4City multi-data map widget, which is designed to integrate heterogeneous Smart City urban data in a 3D map visualization, such as: 3D building models at different LoD, terrain and elevation representation, orthomaps, road networks and traffic flows,

Points of Interest (POI), urban services, dynamic real-time heatmaps, environmental data and information from thousands of sensors distributed across the territory. All these features, together with the capabilities to perform predictive simulations and what-if analysis scenarios, transform the Digital Twin in the Snap4City environment into a fully operational tool, rather than a mere static representation of the city. An example of a Snap4City dashboard showing a multi-data map for 3D Digital Twin representation is shown in Fig. 3.

The graphic rendering in the Snap4City multi-data map widget is based on the Deck.gl framework, which provides the abstraction layer required to handle multiple categories of geographic data and 3D content in a web-based environment. As described in [2], this architecture has been extended with the introduction of elements, such as the *FusionLayer*, which are dedicated to improve the efficiency of data management and visualization for large urban scenarios. The *FusionLayer* is derived from the Deck.gl *TileLayer*, and it is responsible for handling tiled data in a smarter way, allowing for instance to reduce redundant server requests, and limiting client-side computational burden. This is achieved by exploiting client-side cache reuse through *topDownFusion* and *bottomUpFusion* methods that reuse already downloaded data across parent and child tiles during zoom-in and zoom-out operations, respectively. This is particularly relevant for 3D buildings, which require tiled representations and visualization at different zoom levels without providing excessive GPU load or undesirable data duplication. Accordingly, the proposed visualization framework preserves responsiveness even when thousands of buildings and multiple thematic layers are loaded simultaneously.

On top of the above-described architecture, several new interactive features have been introduced for the visualization and management of 3D building models. First, a dynamic loading mechanism has been implemented to replace the previous approach based on static JSON files. In the specific, whenever a request for showing 3D modeled buildings is made in the frontend, a backend script automatically retrieves the server directory structure associated with the currently visible tiles, in order to access the corresponding 3D model. This is done for each visible tile in the viewport. This mechanism ensures that the rendered 3D models are always updated and consistent with the server filesystem structure, since new models or modifications to existing ones become immediately available. This solution makes 3D model management and visualization more flexible and scalable, simplifying the integration of new models and the update of existing ones, without the necessity to manually update configuration files.

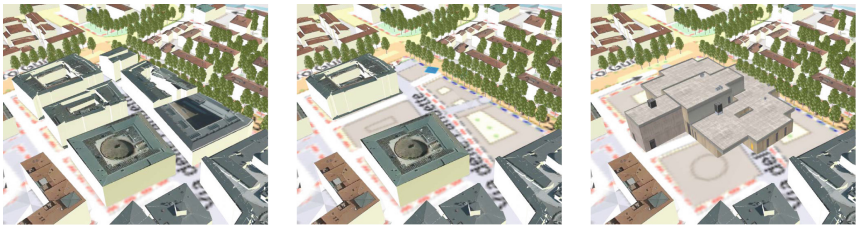
In addition, an interactive building selection and semantic data retrieval from the Knowledge Base have been added as new features: when the user clicks on a building in the 3D view, the selected object is highlighted, and the system queries the Knowledge Base to retrieve the semantic information associated with the selected building previously ingested according to the defined building data model (see Sect. 2 and Fig. 3(b)). In this



Fig. 3. Snap4City Dashboard showing a multi-data map for 3D Digital Twin representation. In (a) full view of the city center (rendering more than 30000 independent building entities) together with IoT sensors, traffic flow information, pollution heatmap. In (b) a close-up view on the Santa Maria del Fiore Cathedral that has been selected to show the related information in the right-sided panel.

way, an informative panel in the multi-data map can show both static and dynamic information related to the specific entity, such as address, building type, number of floors, height, capacity, and also external references (e.g., website, Wikipedia, Wikidata links). Therefore, the 3D model is not treated as a static graphic object, but as the visual counterpart of a semantically indexed urban entity.

Furthermore, the above-described informative panel is also designed to change the visualization model for the selected building at runtime: in addition to showing the metadata, the panel also lists the available 3D models (e.g., GLB/GLTF) for the specific building. When a different model is selected, the default model view is removed from the main scene graph and the new representation is instantiated in a dedicated layer for modified buildings, preserving the consistency between the main building layer and the modified one. As a consequence, the model substitution can be performed interactively without compromising the efficiency of the tiled rendering pipeline. Finally, it is possible to remove one or more buildings from the current scene, which is useful for implementing customized what-if scenarios altering the urban landscape. Actually, by temporarily hiding selected buildings, users can explore alternative urban configurations and visually assess the impact of urban design changes in the surrounding context. Moreover, by exploiting the combination of the model change and remove features, it is possible to introduce in the map new buildings spanning over an area previously covered by different buildings. This allows decision makers to evaluate the impact of renovating full neighborhood not limited to single constructions. An example of this feature is shown in Fig. 4.



(a) (b) (c)

Fig. 4. Example of building removal and full neighborhood modification. In (a) the original neighborhood. In (b) some on the building have been removed. In (c) all the original buildings have been removed leaving space for a new construction. All these changes can be executed directly on the dashboard by the user that select which 3D model visualize.

Therefore, all the above presented features allow the Snap4City dashboards to act as a front-end access point for inspecting semantically indexed entities and interacting with data-driven urban scenarios to monitor and evaluate possible changes of the urban environment.

4 Validation on the SCDT of Florence

To assess the effectiveness and scalability of the proposed framework, a comprehensive validation has been performed in the context of the realization of the SCDT of Florence, Italy. Florence is a medium-sized European city with a historically dense and morphologically complex urban environment, thus representing a challenging case study for urban digital twin applications.

The modelled area encompasses 51,029 buildings. Each building has been independently indexed as a semantic entity in the KM4City knowledge base, following the ingestion pipeline described in Sect. 2, and associated with multiple 3D models. Indeed, two buildings models (with high- and low-resolution roof textures) have been associated with all the buildings, and, when available, additional high quality 3D models and BIM have been considered for specific buildings. The buildings are spatially distributed across 786 tiles, collectively covering approximately 150 km^2, an extent wider than the administrative boundaries of the Florence municipality that includes surrounding areas. Using this scale allows to validate the framework in a concrete and real scenario, not limited to a proof-of-concept experiment.

The two ingestion flows (i.e., for ways and relations) have been executed exploiting a local copy of OSM and 3D models reconstructed from DTM/DSM data. The fault-tolerant design of the pipeline proved essential to efficiently ingest more than 50,000 entities. During ingestion, 169 OSM buildings have been found without an associated 3D model, thus discarded. On the other hand, a number of errors have been logged during the process: 468 errors happened due to temporary service unavailability and have been automatically recovered with the retry procedure. Differently, 84 errors required manual input corrections, after which the ingestion process concluded successfully.

The Snap4City multi-data map dashboard has demonstrated the ability to render the complete Florence digital twin interactively via web-browser. The dashboard is able to load and render thousands of independent building entities simultaneously together with additional data such as IoT sensors, traffic flow information, heatmap, etc. Ingested buildings have been correctly visualized and entity data is efficiently retrieved by clicking the 3D model, thus triggering API calls to perform knowledge base query and obtain entity attributes. Model substitution and removal have been tested thoroughly, evaluating also the possibility of altering entire urban blocks, thus confirming the effective operability of the what-if scenario mechanism.

These results support the conclusion that the proposed framework can be adopted in real scenarios to realize SCDT of medium and large municipalities.

5 Conclusions

In this paper a framework for the semantic ingestion and representation of 3D building entities within a Smart City Digital Twin has been presented. Exploiting the KM4City ontology and specific data models, the system is able to index buildings as semantically enriched knowledge base entities related to multiple 3D models, moving forward with respect to static 3D urban representations. In order to accomplish such a result, a scalable and fault-tolerant ingestion pipeline capable of reconciling OpenStreetMap metadata with 3D models has been developed using the processing logic built-in the Snap4City platform. In addition, dashboards within the Snap4City platform provide a high-performance 3D visualization environment that supports a wide range of dynamic interactions to allow users to select buildings and retrieve semantic data, substitute 3D models at runtime, remove structures, and redesign complete areas to simulate what-if scenarios. The framework has been validated on the case study of Florence, Italy, that required the management of over 50,000 independent building entities covering an

area of approximately 150 km^2. Therefore, the proposed solution, providing semantic representations of building and interactive widely accessible web-based 3D interfaces, realizes a comprehensive tool supporting decision-makers for the operational monitoring and strategic planning of smart cities, and allows common citizens to navigate through their city and be involved in the urban development.

Acknowledgments. This work has been carried out in the context of the Interreg Euro-MED TOURISMO project "TOURism Innovative and Sustainable Management of flOws", co-funded by the European Union (Euro-MED0200900).

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

References

1. Taylor, J., Bennett, G., Mohammadi, N.: Engineering smarter cities with Smart City digital twins. J. Manag. Eng. **37**(6) (2021). https://doi.org/10.1061/(asce)me.1943-5479.0000974
2. Adreani, L., Bellini, P., Fanfani, M., Nesi, P., Pantaleo, G.: Smart City digital twin framework for real-time multi-data integration and wide public distribution. IEEE Access. **12**, 76277–76303 (2024). https://doi.org/10.1109/ACCESS.2024.3406795
3. Nesi, P., Garau, C., Zamperlin, P.: Ottimizzazione delle Città Intelligenti: L'Utilizzo Dei gemelli Digitali e dell'Intelligenza Artificiale nei Processi Decisionali Urbani. CONTESTI. **2**, 1–13 (2025)
4. Cheng, R., Hou, L., Xu, S.: A review of digital twin applications in civil and infrastructure emergency management. Buildings. **13**(5), 1143 (2023). https://doi.org/10.3390/buildings13051143
5. Cureton, P., Hartley, E.: City information models (CIMs) as precursors for urban digital twins (UDTs): a case study of Lancaster. Front. Built Environ. **9** (2023). https://doi.org/10.3389/fbuil.2023.1048510
6. KM4City, the DISIT knowledge model for city and mobility. Available at http://wlode.disit.org/WLODE/extract?url=http://www.disit.org/km4city/schema. Accessed 3 Mar 2026
7. Bellini, P., Benigni, M., Billero, R., Nesi, P., Rauch, N.: Km4City ontology building vs data harvesting and cleaning for smart-city services. J. Vis. Lang. Comput. **25**(6), 827–839 (2014). https://doi.org/10.1016/j.jvlc.2014.10.023
8. Bellini, P., Collini, E., Fanfani, M., Nesi, P., Panconi, C.: Efficient and scalable semantic data ingestion for Smart City digital twin platforms. In: 2025 IEEE 11th International Conference on Big Data Computing Service and Machine Learning Applications (BigDataService), Tucson, AZ, USA, pp. 44–51 (2025). https://doi.org/10.1109/BigDataService65758.2025.00012
9. Bilotta, S., Nesi, P.: Traffic flow reconstruction by solving indeterminacy on traffic distribution at junctions. Futur. Gener. Comput. Syst. **114**, 649–660., ISSN 0167-739X, (2021). https://doi.org/10.1016/j.future.2020.08.017
10. Bilotta, S., Ipsaro Palesi, L.A., Nesi, P.: Exploiting open data for CO2 estimation via artificial intelligence and eXplainable AI, 128598., ISSN 0957-4174. Expert Syst. Appl. **291** (2025). https://doi.org/10.1016/j.eswa.2025.128598
11. Fanfani, M., Ipsaro Palesi, L.A., Nesi, P.: Data-driven mobility demand-versus-offer matching by dynamic integration of public and sharing services. IEEE Access. **14**, 23166–23193 (2026). https://doi.org/10.1109/ACCESS.2026.3661850

12. Bellini, P., Fanfani, M., Nesi, P., Pantaleo, G.: Snap4City dashboard manager: a tool for creating and distributing complex and interactive dashboards with no or low coding, 101729., ISSN 2352-7110. SoftwareX. **26** (2024). https://doi.org/10.1016/j.softx.2024.101729
13. Adreani, L. et al.: Smart City scenario editor for general what-if analysis. Sensors. **24**, 2225 (2024). https://doi.org/10.3390/s24072225
14. De Nicola, A., Villani, M.L.: Smart City ontologies and their applications: a systematic literature review. Sustainability. **10**(10), 5578 (2021). https://doi.org/10.3390/su13105578
15. Bittner, T., Donnelly, M., Winter, S.: Ontology and semantic interoperability. In: Large-Scale 3D Data Integration. CRC Press (2005)
16. Fernandez, S., Hadfi, R., Ito, T., Marsa-Maestre, I., Velasco, J.R.: Ontology-based architecture for intelligent transportation systems using a traffic sensor network. Sensors. **16**(8), 1287 (2016)
17. Shi, J., Pan, Z., Jiang, L., Zhai, X.: An ontology-based methodology to establish city information model of digital twin city by merging BIM, GIS and IoT. Adv. Eng. Inform. **57**, 102114 (2023). https://doi.org/10.1016/j.aei.2023.102114
18. Ju, J., Liu, L., Feng, Y.: Citizen-centered big data analysis-driven governance intelligence framework for smart cities. Telecommun. Policy. **42**(10), 881–896 (2018). https://doi.org/10.1016/j.telpol.2018.01.003
19. D'Aniello, G., Gaeta, M., Orciuoli, F., Sansonetti, G., Sorgente, F.: Knowledge-based Smart City service system. Electronics. **9**(6), 965 (2020). https://doi.org/10.3390/electronics9060965
20. Sobral, T., Galvão, T., Borges, J.: Knowledge-assisted visualization of multi-level origin-destination flows using ontologies. IEEE Trans. Intell. Transp. Syst. **22**(4), 2168–2177 (2021). https://doi.org/10.1109/TITS.2021.3056228
21. s4city: SAREF extention for smart city. Available at https://saref.etsi.org/saref4city/v1.1.2/. Accessed 3 Mar 2026
22. SAREF: The smart applications REFerence onotology. Available at https://saref.etsi.org/core/v3.2.1/. Accessed 3 Mar 2026
23. Building Topology Ontology. Available at https://w3c-lbd-cg.github.io/bot/. Accessed 3 Mar 2026
24. Gröger, G., Plümer, L.: CityGML–interoperable semantic 3D city models. ISPRS J. Photogramm. Remote Sens. **71**, 12–33 (2012). https://doi.org/10.1016/j.isprsjprs.2012.04.004
25. Adreani, L., Bellini, P., Colombo, C., et al.: Implementing integrated digital twin modelling and representation into the Snap4City platform for smart city solutions. Multimed. Tools Appl. **83**, 37121–37146 (2024). https://doi.org/10.1007/s11042-023-16838-0
26. Fanfani, M., Ipsaro Palesi, L.A., Nesi, P.: Microservices' libraries enabling server-side business logic visual programming for digital twins, 101805., ISSN 2352-7110. SoftwareX. **27** (2024). https://doi.org/10.1016/j.softx.2024.101805
27. Node RED framework. Available at https://nodered.org/. Accessed 3 Mar 2026

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

Integrating GIS-Based Spatial Analysis and Structural Modeling for the Adaptive Reuse of the Santa Chiara Dam

Enrica Vecchi(✉) and Marco Zucca

Department of Civil, Environmental Engineering and Architecture, University of Cagliari, Cagliari 09123, Italy
{enrica.vecchi,marco.zucca2}@unica.it

Abstract. Adaptive reuse of historic hydraulic infrastructures requires precise geometric reconstruction to enable reliable structural assessment and design-oriented evaluations. This study presents an integrated workflow combining GIS-based spatial analysis and structural modeling, applied to the Santa Chiara Dam, a decommissioned early 20th-century gravity–multiple-arch dam located on the Tirso River (Sardinia, Italy).

A key challenge addressed in this work is the lack of direct bathymetric information for the partially submerged reservoir basin. Existing high-resolution LiDAR-derived Digital Terrain Models (DTMs, 1 m resolution) provided by the Sardinian Geoportal capture only the water surface, leaving the underlying morphology undefined. For partially submerged dams, even an approximate reconstruction of the reservoir basin is essential, not only for documenting the structure itself but also for defining the spatial context required for reliable structural assessment and load distribution analysis. To overcome this limitation, a dedicated integrated workflow was developed, combining GIS-based spatial analysis with structural modeling. The workflow includes: (i) extraction and clipping of external terrain morphology; (ii) slope analysis of surrounding banks; (iii) geometric reconstruction of transition zones through transversal profiles and linear gradient assumptions; (iv) generation of a continuous terrain model via TIN interpolation and raster merging; and (v) non-linear finite element analysis of the dam structure using Midas FEA software.

Results indicate an overall satisfactory structural response, with limited displacements and predominantly compressive stress states, while highlighting localized tensile stress concentrations at arch–buttress interfaces.

The proposed workflow demonstrates the value of integrating GIS-based spatial analysis with structural modeling for heritage hydraulic structures, facilitating both accurate structural evaluation and sustainable adaptive reuse strategies. The methodology is transferable to similar historical infrastructures where incomplete topographic or bathymetric data limit conventional assessment approaches.

Keywords: Adaptive Reuse · Historic Dams · GIS · Spatial Analysis · Structural Modeling · Digital Terrain Model · Finite Element Analysis · Cultural Heritage · Sardinia

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 37–53, 2027.
https://doi.org/10.1007/978-3-032-30539-8_3

1 Introduction

The rehabilitation and reuse of existing infrastructure have become one of the most important and timely research topics in civil and environmental engineering, particularly in the context of aging assets and the growing demand for sustainable construction strategies. Over the past decades, numerous projects have therefore focused on the rehabilitation of large-scale infrastructures of significant historical and cultural value, with the aim of extending their service life while preserving their structural and architectural identity [1–3]. Among these infrastructures, dams stand out because of their strategic importance, structural complexity, and the potential consequences associated with their deterioration or abandonment. Indeed, even when no longer operational, many of these structures constitute valuable elements of industrial heritage, providing important evidence of past engineering practices and technological development. For this reason, the assessment, preservation, and possible rehabilitation of such infrastructures have become key topics in both structural and conservation engineering [4, 5].

Historical dams are often characterized by complex geometries that are difficult to reconstruct, particularly in portions that remain permanently submerged, where direct inspection and geometric documentation are typically limited. Furthermore, the wetdry cycles affecting exposed areas may lead to significant material degradation, influencing both primary and secondary structural elements. Different approaches have been developed to address these challenges, integrating geometric reconstruction with advanced structural analysis methods.

Previous studies have explored different modeling strategies. In [6], the structural performance of a large arch dam was evaluated under both static and dynamic loads, with particular attention to joint movements. Fedele et al. [7] proposed a new approach to assess the influence of the alkali-silica reaction and the presence of diffused microcracking on structural performance of existing concrete dams, depending on the age of the structure. A numerical model, based on the Discrete Element Method (DEM), was proposed in [8] to investigate failure mechanisms involving the dam body and the damrock interface. In particular, dedicated tools for the analysis of masonry gravity dams have been developed, using the Guilhofrei Dam as a case study. Nguyen-Tuan et al. [9] developed a coupled thermo-hydro-mechanical model for masonry dams optimized through parallel computation. Robles and Pena [10] analyzed the influence of masonry mechanical properties on the structural behavior of masonry dams. Results of parametric analyses identified tensile strength as a key parameter governing structural response. The study reported in [11] investigates the condition and potential reuse of stone masonry from the Kada Dam after 46 years of service, highlighting that the blocks preserve sufficient residual strength to be employed in new gravity dams' construction.

In this study, the structural behavior of the Santa Chiara Dam on the Tirso River (Sardinia) is investigated starting from a GIS-based reconstruction of the riverbed geometry. A dedicated workflow integrating GIS-based spatial analysis, supported by custom python routines, with structural modeling is proposed. Completed in 1924 based on the design by Luigi Kambo and Angelo Omodeo, the dam played a key role in the economic and social development of the region and is currently the subject of a rehabilitation program.

The paper is organized as follows. Section 2 describes the dam from a historical, structural, and geological perspective. Section 3 presents the methodology used for the reconstruction of the Digital Terrain Model (DTM). Section 4 describes the Finite Element Modeling (FEM) and discusses the main results of the non-linear static analysis, considering the safety requirements imposed by the rehabilitation project. Finally, Sect. 5 summarizes the main findings and outlines directions for future developments.

This integrated approach enables the definition of a consistent geometric framework linking terrain morphology, hydraulic context, and structural configuration of the dam, supporting both structural assessment and the sustainable management of historical hydraulic infrastructures.

2 Santa Chiara Dam

2.1 Brief Historical Notes

The Santa Chiara Dam is located in central Sardinia (Italy), in the Province of Oristano, within the municipality of Ulà Tirso, along the Tirso River. Construction took place between 1918 and 1924, although initial design studies for the dam and its artificial reservoir date back to 1912 (Fig. 1). The project was conceived by the engineer Angelo Omodeo and executed by the company *"Imprese Idrauliche ed Elettriche"*, with the aim of flood regulation along the Tirso River, hydroelectric power generation, and irrigation to the Campidano plain. After the dam was completed, the upstream artificial reservoir was named Lake Omodeo to honor the project designer.



Fig. 1. Santa Chiara dam today.

The artificial reservoir was initially realized as a water storage system intended to support irrigation, but the project was subsequently revised and adapted to include hydroelectric power generation. This required increasing the reservoir capacity and relocating the dam to exploit a higher hydraulic head and improve the efficiency of the power plant.

Construction works of the gravity dam began in 1918 under engineer Luigi Kambo's supervision and continued until 1924. During foundation excavation, a tuffaceous substratum was discovered, which was neither sufficiently compact nor mechanically resistant to support the stresses predicted at the dam's base by the original design. This condition would have required substantially larger concrete volumes, making the original gravity dam design economically unfeasible. In agreement with Kambo, Omodeo and Giulio Dolcetta proposed significant modifications, ultimately constructing a multiple-arch dam 260 m long and 70 m high (Fig. 2).

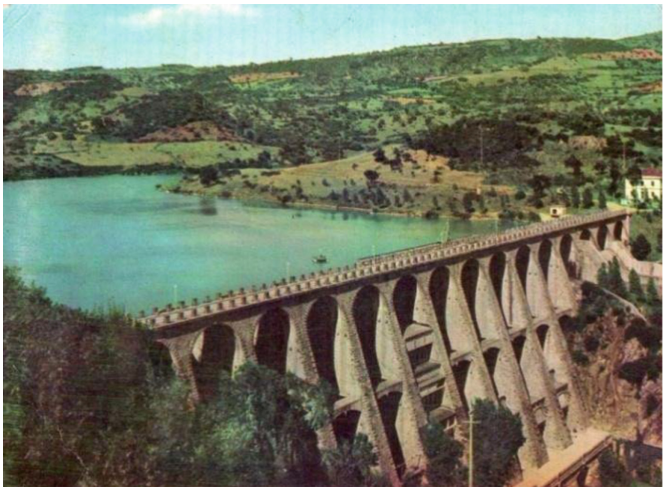


Fig. 2. Santa Chiara Dam during its first year of operation [12].

The project was completed in 1924 and officially inaugurated on 28 April of the same year. The dam remained in operation until 1968, when structural issues in the buttresses led to operational limitations. Although a replacement dam downstream was considered to increase water storage capacity, Santa Chiara was formally decommissioned only in 1997 and fully taken out of service by 2000.

2.2 Structural and Geological Configuration

The dam is a composite masonry-concrete structure, consisting of 17 spans with a clear span of 15 m each. The maximum height from the foundation is 61 m, reaching an elevation of 109 m above sea level. The structural layout follows the typical scheme of gravity dams.

The buttresses are reinforced by a four series of semicircular arches with a span of 12.5 m, above which a 6 m-wide roadway runs at a constant elevation of 112 m a.s.l. Figure 3 illustrates the plan view of the dam, indicating the main structural elements.

The buttresses and arches are constructed of trachyte masonry with cement mortar joints, whereas the downstream shells are made of concrete. Original design documentation indicates a mortar composition with 300 kg/m^3 cement, a compressive strength

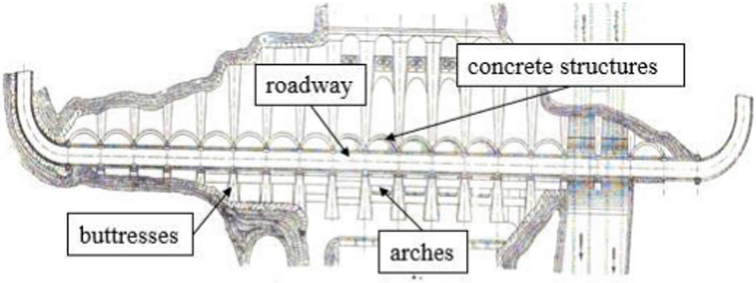


Fig. 3. Dam plan view.

of 27 MPa, and a tensile strength of 0.35 MPa after 180 days. After 28 days, the mortar exhibits a compressive strength of 13.5 MPa and a tensile strength of 0.15 MPa, while the concrete can be considered equivalent to a C25/30 strength class with a unit weight equal to 25 kN/m^3. Trachyte masonry shows a compressive strength of 3 MPa, tensile strength of 0.3 MPa, and a unit weight of 22 kN/m^3, as reported in [13]. Table 1 summarizes the key mechanical parameters of materials used in the main structural elements.

Table 1. Mechanical properties of main structural elements

Element	Material	Elastic Modulus [MPa]	Compressive strength [MPa]	Tensile strength [MPa]
Buttresses	trachyte trachyte	2400	3	0.3
Arches Shells	concrete	2400	3	0.3
		29961	30	3

From a geological and lithological perspective, the study area is characterized by Cenozoic volcanic rocks, mainly Upper Oligocene to Middle Miocene volcanites, locally associated with a late Hercynian magmatic complex of intrusive and volcanic origin. The foundation bedrock is primarily constituted by compact trachyte flows, overlain in places by continental and marine sediments from the Upper Paleocene to Pliocene. The geological setting is further influenced by a Cenozoic post-Hercynian fault system that crosses a significant portion of the reservoir approximately perpendicular to its main axis, representing a major structural discontinuity within the rock mass.

The surface stratigraphy includes vesicular and compact basalts, trachytic tuffs, and compact trachytes, which constitute the primary volcanic units of the area. The wider catchment basin is mainly composed of granites and crystalline schists, which are particularly prone to weathering and thus contribute to the formation of a weathered mantle and localized alluvial deposits. In the western sector, ancient trachytes form the basal substratum, while basaltic lava flows create plateau structures occasionally overlaying them, sometimes associated with volcanic tuff deposits. Alluvial areas on the plateaus mainly consist of clayey-sandy soils with rounded quartz pebbles, whose size and abundance progressively decrease toward the coastal plain.

3 Gis-Based Spatial Analysis

This chapter describes the methodology adopted to reconstruct a simplified model of the riverbed in the area surrounding the Santa Chiara dam. The reconstruction is based on the integration of publicly available geospatial datasets, historical drawings, technical documentation, geometric assumptions, and information derived from the literature, all combined within a GIS environment to produce a consistent and reproducible representation of the reservoir basin. The procedure was implemented using the opensource software QGIS (version 3.34) [14] supported by custom python scripting and is fully reproducible and transferable to similar case studies characterized by incomplete geomorphological data. The following sections first present the available UAV photogrammetric dataset (Sect. 3.1), then discuss the limitations of publicly available Digital Elevation Models in representing submerged terrain (Sect. 3.2) and finally describe the simplified reconstruction of the riverbed geometry based on GIS analysis and geometric assumptions (Sect. 3.3).

3.1 UAV Photogrammetric Model of the Emerged Structure

A recent UAV-based photogrammetric model of the dam and its surrounding environment was available for this study, providing a detailed three-dimensional representation of the structure and adjacent terrain. The dataset was acquired using a DJI Mavic 3 platform equipped with an onboard camera featuring a nominal focal length of 4.4 mm (equivalent to 24 in 35 mm format). A total of 939 images were collected through combined nadir and oblique flight configurations.

The images were processed using a Structure-from-Motion workflow in Agisoft Metashape (version 2.2.2) [15]. Georeferencing was performed using 26 Ground Control Points (GCPs) surveyed with GNSS in NRTK mode, achieving an accuracy on the order of 3–4 cm in the Northing and Easting and about 5 cm in the height. The main products include a dense point cloud (Fig. 4) and a high-resolution orthophoto.

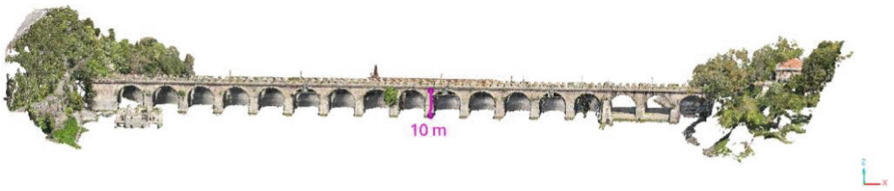


Fig. 4. Dense point cloud derived from UAV photogrammetry (SfM), georeferenced using GNSS-surveyed GCPs. The model provides a detailed 3D representation of the visible portions of the dam and surrounding terrain.

Although the photogrammetric dataset provides a detailed three-dimensional representation of the visible portions of the dam and surrounding terrain, it does not capture reliable elevation information for the submerged riverbed due to the presence of water

surfaces and their significant seasonal variability [16]. This limitation is inherent to optical survey techniques, as water bodies prevent the reconstruction of underlying terrain features using standard SfM approaches [17].

Within the proposed workflow, the UAV model is therefore primarily used to support the geometric characterization of the exposed structure and the spatial interpretation of the study area, serving as a complementary dataset.

3.2 Limitations of Publicly Available DTMs

A reliable representation of the riverbed is essential for the structural assessment of partially submerged hydraulic infrastructures such as the Santa Chiara Dam. Indeed, since the present study relies on historical drawings and technical documentation, reconstructing the submerged riverbed is necessary to correctly define the dam geometry and its interaction with the surrounding terrain. This aspect is particularly relevant in heritage engineering applications, where incomplete or indirect data must be integrated to obtain consistent structural models [18]. Moreover, the reconstructed basin is crucial for setting boundary conditions and loading scenarios in the finite element model, enabling a realistic representation of hydrostatic pressure along the dam body.

However, none of the publicly available Digital Elevation Model datasets for the study area provide direct information on the morphology of the riverbed beneath the water surface, as they only capture the water surface at the time of acquisition. It is worth noting that this limitation is common to many open-access geospatial datasets, including regional geoportals and national mapping services, which are typically optimized to represent emerged terrain rather than submerged environments. This issue is also inherent to DTMs derived from standard airborne LiDAR surveys, as laser pulses are unable to penetrate water unless specifically designed for bathymetric measurements [19, 20]. On the other hand, while dedicated bathymetric datasets and hydrographic surveys are available for major rivers and large water bodies (e.g., USGS bathymetric surveys [21], NOAA hydrographic survey data [22], and GEBCO global bathymetry [23])), they are typically absent for smaller reservoirs such as the one considered in this study [24]. In such cases, information on water bodies is often limited to planimetric data (e.g., shoreline digitization), with no associated depth information, unless dedicated field surveys are conducted.

Within this context of limited data availability, the primary geospatial information for the Santa Chiara Dam area was obtained from the Geoportal of the Autonomous Region of Sardinia [25]. Among the available datasets, the high-resolution Digital Terrain Model (DTM) produced from airborne LiDAR surveys in 2010 by the Italian Ministry for the Environment was chosen for this study. The DTM belongs to the "DTM River Bands" collection, with a spatial resolution of 1 m, and is referenced to the ETRF2000 / UTM Zone 32 N coordinate system (EPSG:6707), the official national reference system in Sardinia.

As mentioned, this dataset, acquired for topographic and hydraulic risk assessment purposes, represents water-covered areas as flat surfaces corresponding to the water level, without including bathymetric information. This limitation is illustrated in Fig. 5, which shows the regional DTM (a) together with an elevation profile extracted across the reservoir using the QGIS Profile Tool plugin (b). While the surrounding terrain

exhibits realistic topographic variability, the profile within the reservoir remains nearly constant, clearly indicating that the model represents the water surface rather than the actual riverbed.

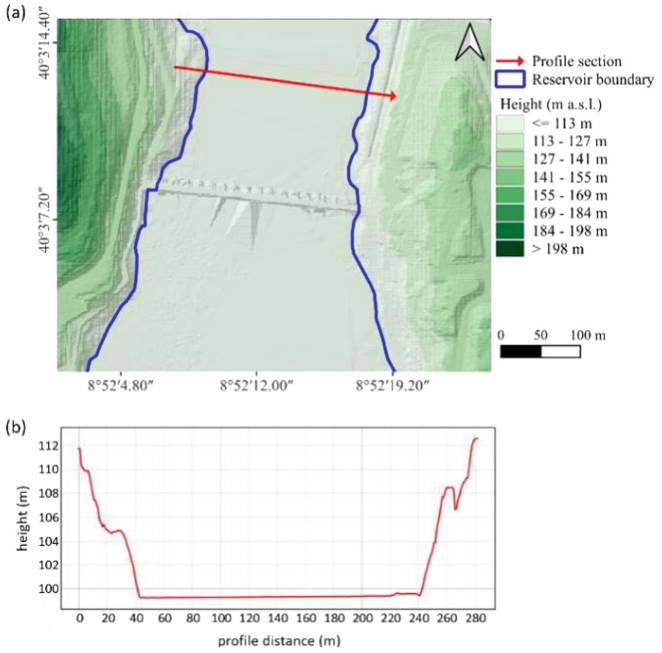


Fig. 5. Regional LiDAR-derived DTM of the study area, showing the location of the extracted profile and the reservoir boundary; (b) corresponding elevation profile extracted using the Qgis Profile Tool plugin. Elevations are expressed as orthometric heights (m a.s.l.), and coordinates are referenced to the WGS84 system (EPSG:4326).

To address this issue, which prevent the dataset from being directly used due to misleading heights, water-covered portions of the DTM were excluded by manually digitizing the reservoir boundary from recent satellite imagery. This allowed a clear separation between areas reliably captured by LiDAR and sectors affected by water-surface representation. The resulting clipped DTM was then used to derive a slope map, which provided an initial characterization of the terrain and informed the geometric assumptions for reconstructing the submerged basin (Fig. 6).

3.3 Simplified Reconstruction of the Riverbed

Based on the pre-processed terrain data described in the previous section, a simplified reconstruction of the reservoir basin was developed using geometric assumptions and other available information. The adopted configuration was intended as an approximate representation of the submerged morphology rather than a fully representative geomorphological model of the actual terrain. This approach was considered a reasonable

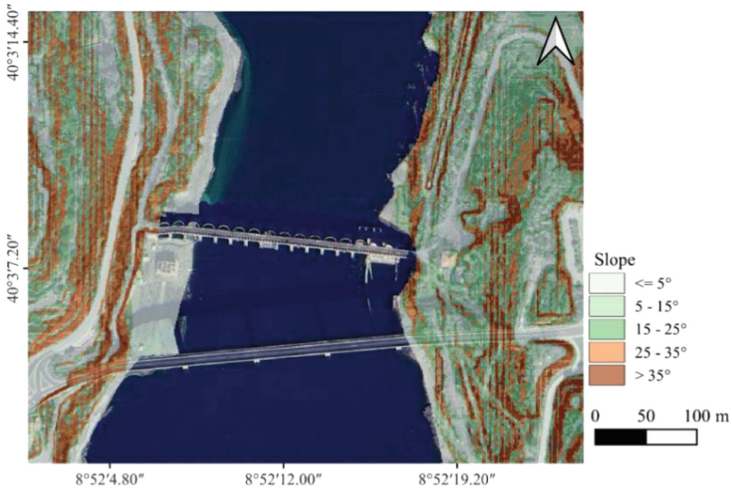


Fig. 6. Slope map derived from the LiDAR-based DTM provided by the Sardegna Geoportal. Grid coordinates are aligned to the WGS84 reference system (EPSG:4326).

compromise between realism, data availability, and methodological feasibility for the purposes of the present study.

The deepest portion of the reservoir basin was represented as a uniform-elevation surface, corresponding to the central flat portion of the riverbed. According to the original technical documentation, this elevation was set to 50 m above mean sea level [12]. The central polygon was rasterized within the GIS environment, assigning the constant elevation value to each cell. The resulting raster was generated at 1 m spatial resolution and in the same coordinate reference system as the original LiDAR-derived DTM, ensuring seamless integration with the surrounding terrain. This simplified flat-bottom surface serves as the foundation for defining the lateral transition zones connecting the riverbed to adjacent terrain (Fig. 7a).

The lateral sections of the basin were reconstructed using a simplified geometric approach. The cross-section was idealized as an isosceles trapezoid, with linear slopes bridging the elevation difference between the riverbed and the surrounding terrain. The definition of representative slope values was therefore a key step in ensuring a geometrically consistent and physically plausible model.

Preliminary analyses included the extraction of slope gradients along multiple transects derived from the DTM. However, to reduce the influence of local topographic irregularities and ensure consistency with historical design information, a simplified geometric approach based on known reservoir dimensions was ultimately adopted.

According to this approach, the river's total width (approximately 250 m) and the width of the central flat-bottom region (100 m) define two symmetrical transition zone of 75 m on each side. The elevation difference between the surrounding terrain (≈99 m a.s.l.) and the riverbed (50 m a.s.l.) results in a vertical drop of 49 m. Assuming a linear profile, the slope of the riverbanks was defined as the ratio between vertical and horizontal distance, resulting in an inclination of approximately 33° (about 65%).

This value was considered representative of the overall basin geometry, as it provides a robust approximation that is less affected by local topographic noise compared to purely data-driven slope estimations [26]. Based on this assumption, the eastern and western transition zones were delineated by offsetting the river boundary line by a uniform distance of 75 m, resulting in two distinct polygons representing the sloping riverbanks.

A continuous elevation surface for the transition zones was generated using a transect-based workflow developed through custom Python scripts. Points were initially sampled along the river boundary at 1 m intervals, and each point was assigned the elevation of the nearest terrain DTM pixel, representing the interface between emerged terrain and water-covered areas.

From these boundary points, transects extending toward the interior of the reservoir were generated. Along each transect, additional points were distributed at 0.5 m intervals, and their elevations were linearly interpolated from the terrain level down to the riverbed elevation, following the predefined slope.

The resulting sets of points, representing distributed elevation values along the transition zones, were interpolated using a Triangulated Irregular Network (TIN) algorithm, to generate continuous raster surfaces at 1 m resolution (compatible with the original LiDAR-derived terrain model). Figure 7 illustrates the intermediate raster layers used in the reconstruction process, including the constant-depth central riverbed (a) and the interpolated transition surfaces derived from the transect-based procedure on the eastern and western sides (b).

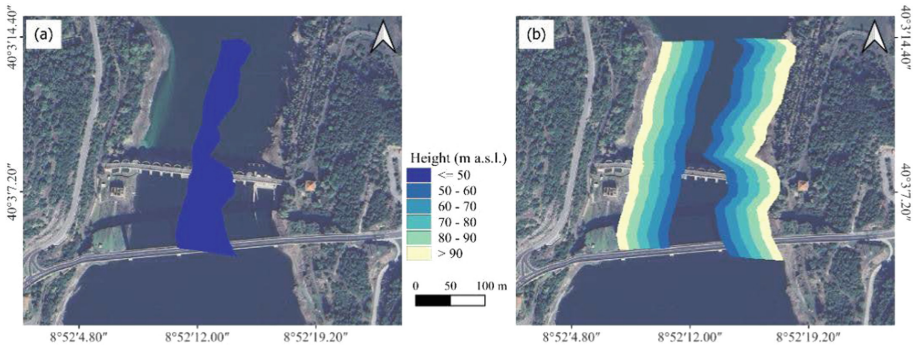


Fig. 7. Intermediate raster layers used for the reconstruction of the reservoir basin: (a) constantdepth central riverbed; (b) interpolated transition surfaces for the eastern and western banks generated from transect-based elevation points. Elevations are expressed as orthometric heights in meters above sea level (m a.s.l.), and grid coordinates are aligned to the WGS84 reference system (EPSG:4326).

The raster datasets representing the central riverbed and the two transition zones were then spatially aligned to ensure identical grid size, origin, and coordinate reference system. Alignment was performed using GDAL tools, enforcing uniform 1 m resolution and consistent grid anchoring. Once aligned, the rasters were merged into a single continuous DTM representing the reconstructed basin. NoData values were carefully handled,

and local interpolation was applied where necessary to eliminate discontinuities and ensure smooth transitions between surfaces.

The final reconstructed DTM provides a continuous and coherent representation of the basin, suitable for both geometric analysis and integration into structural numerical modeling workflows. To facilitate this integration, the DTM was further processed within the GIS environment to extract contour lines describing the riverbed geometry, as shown in Fig. 8.

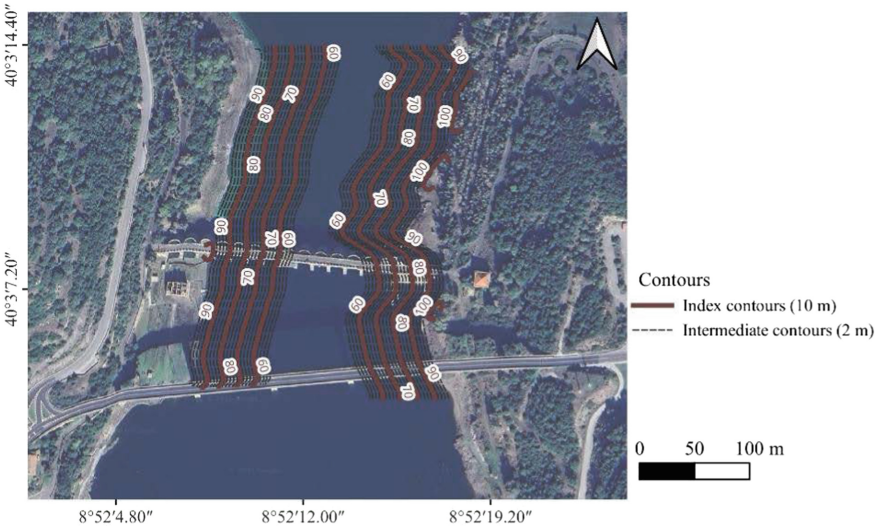


Fig. 8. Contour lines generated from the reconstructed DTM. Elevations are given as orthometric heights in meters above sea level (m a.s.l.), and grid coordinates are aligned to the WGS84 reference system (EPSG:4326).

4 Structural Modelling and Numerical Analysis

4.1 Geometry Reconstruction

The geometric reconstruction of the dam was achieved by integrating data derived from the UAV photogrammetric survey with the available historical documentation. These sources together provided the basis for generating a 3D model of the Santa Chiara Dam within a commercial modelling environment [27]. In particular, the main structural components (buttresses, arches, shells, and slabs) were represented using solid elements, allowing the complex geometric configuration typical of this type of structure to be reproduced with adequate accuracy. Since the initial geometric reconstruction assumed uniform buttress heights, the actual above-ground elevations of the vertical elements were refined by intersecting the structural model with the reconstructed Digital Terrain Model (DTM) developed in the previous phase. The resulting terrain-structure interaction was then resolved through a Boolean subtraction operation, enabling the definition of the

final geometry of the exposed structural components (Fig. 9). This procedure ensured geometric consistency between the dam model and the surrounding terrain.

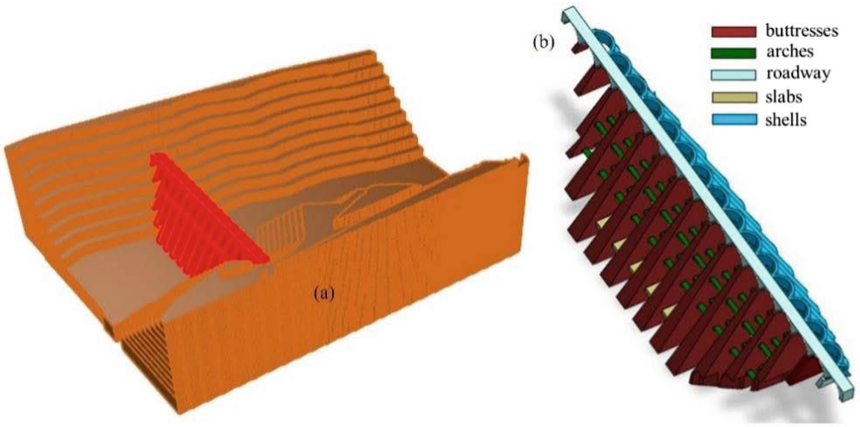


Fig. 9. (a) DTM and (b) geometry reconstruction of the dam.

Starting from this reconstructed geometry, a 3D finite element model (FEM) was implemented using Midas FEA software [28]. The mesh consists of 1264789 tetrahedral solid elements, each with a maximum size limited to 1 m (Fig. 10), ensuring an adequate level of detail for structural analysis.

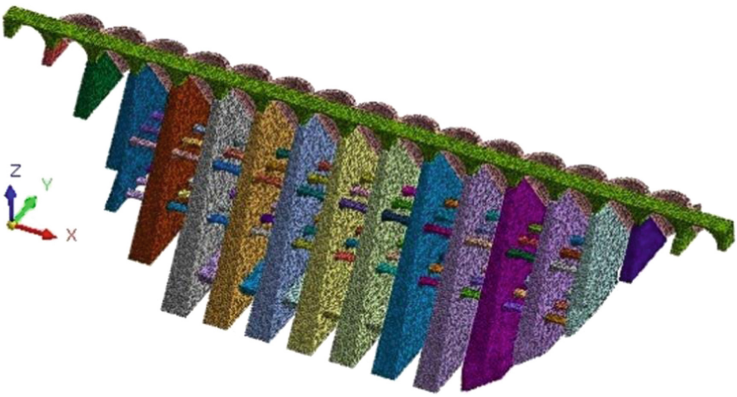


Fig. 10. Finite element model of the dam.

Boundary conditions were defined by applying fixed constraints at the base of the buttresses and along the soil–structure interfaces, assuming a perfectly restrained condition. Three types of loads were considered: (i) self-weight of the structure (automatically computed by the software), (ii) a crowd load applied on the roadway, equal to 4 kN/m^2 [29], and (iii) a horizontal load due to the presence of the water.

To account for the nonlinear behavior of both masonry and concrete, a smeared cracking (CSC) constitutive model was adopted for all materials [30], using the mechanical parameters reported in Table 1.

4.2 Numerical Analysis

The structural response of the dam was investigated through a non-linear static construction-stage analysis. The numerical procedure was implemented in two main phases. In the first phase, the structural model was generated by activating the elements and applying the self-weight together with the hydrostatic pressure of the reservoir, assumed to act up to two-thirds of the total height of the dam. In the second phase, the crowd load acting on the roadway was introduced through a load-stepping procedure, with sequential increments of 10% until the design load of 4 kN/m^2 specified in the rehabilitation project was reached. The resulting vertical displacements field at the end of the second phase is shown in Fig. 11.

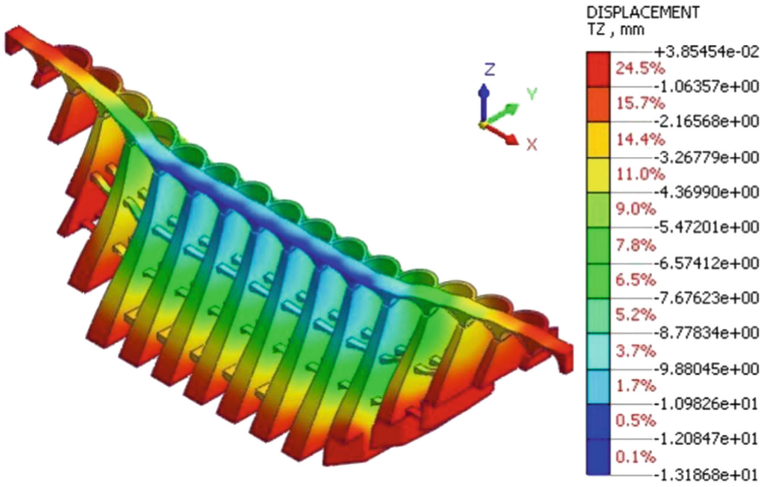


Fig. 11. Vertical displacements at the end of the second phase

The analysis indicates a maximum vertical displacement of approximately 1.32 cm (highlighted in blue), occurring at the central arch span. The displacement distribution is nearly symmetric with respect to the dam centerline, reflecting the regularity of the structural configuration and loading condition. The magnitude of the displacement remains within acceptable serviceability requirements, defined in the dam rehabilitation project, considering the overall length of the structure of approximately 257 m.

The corresponding stress fields, in terms of maximum tensile and compressive stresses, are reported in Figs. 7 and 8, respectively. The numerical results show that the stress field is largely consistent with the material strength capacities, with only limited localized regions exhibiting stress concentrations higher than the admissible limits. The maximum compressive stress is approximately 1 MPa, which is significantly lower

than the characteristic compressive strength of both masonry and concrete. This condition is primarily due to the dominance of the self-weight in the overall load combination, which largely exceeds the contribution of the crowd load (Figs. 12, 13 and 14).

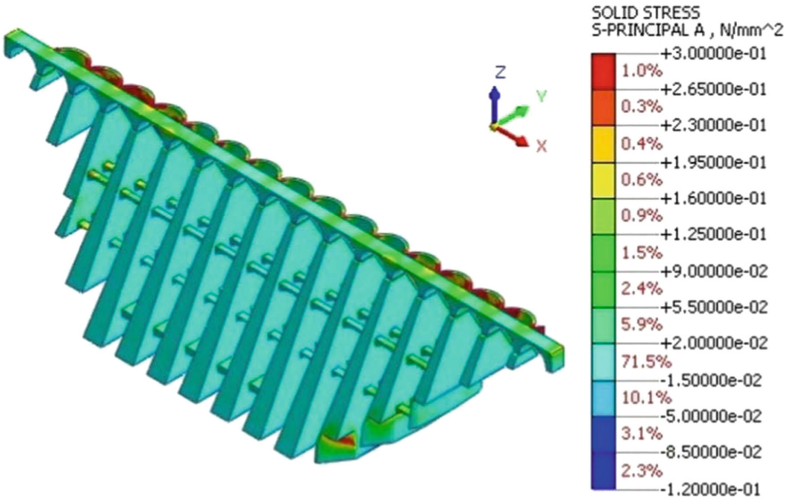


Fig. 12. Maximum tensile stress.

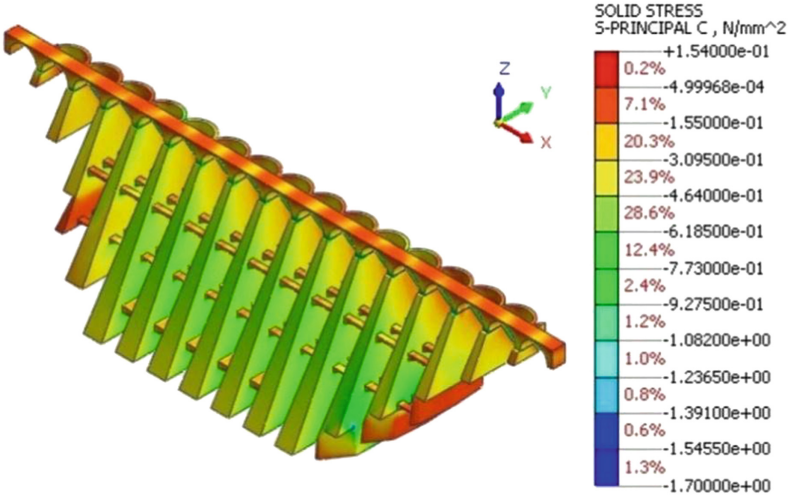


Fig. 13. Maximum compressive stress

Regarding tensile stresses, most areas of the dam remain within the tensile resistance of the materials. However, a few localized regions exhibit tensile stresses exceeding the masonry tensile strength, primarily concentrated in the upper portions of the structure.

This distribution is consistent with the observed structural condition, where cracking is mainly localized in these areas. Figure 9 illustrates the predicted cracking pattern in the masonry regions affected by tensile stress exceedance.

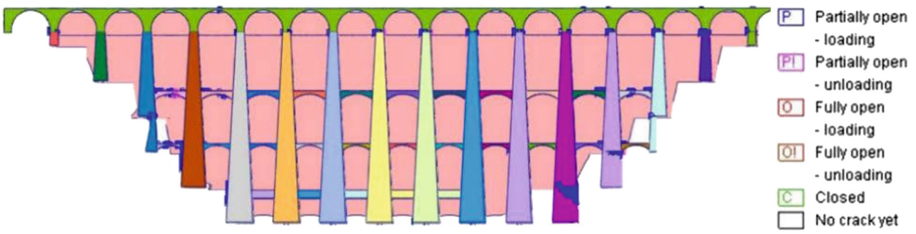


Fig. 14. Cracks distribution

The identified cracking patterns are limited to small, localized regions and are consistent with the principal tensile stress distribution, thus not posing significant concerns for the proposed adaptive reuse of the dam.

5 Conclusions

This study presents an integrated workflow combining GIS-based spatial analysis and structural modeling for the assessment and potential adaptive reuse of the Santa Chiara dam, a historical multiple-arch hydraulic structure.

A key challenge addressed in this work was the incomplete geometric characterization of the reservoir basin, particularly in submerged areas where direct bathymetric data were not available. To overcome this limitation, a simplified reconstruction methodology was developed, based on GIS spatial analysis, using publicly available DTMs, historical documentation, and geometric assumptions. The approach enabled the generation of a continuous and coherent model of the reservoir basin, suitable for subsequent numerical analyses, with the GIS environment acting as a bridge between largescale spatial data and structural modeling, ensuring a consistent transfer of geometrical information from the territorial to the structural scale.

The reconstructed terrain model was therefore successfully integrated into the structural modelling workflow, allowing the reliable definition of boundary conditions and load distributions. Indeed, unlike conventional approaches, where hydrostatic loads are often idealized or simplified, the adopted methodology enabled the explicit definition of the water domain.

The resulting finite element analysis highlighted an overall satisfactory structural behaviour of the dam under the loading conditions imposed by the rehabilitation project. Compressive stress (about 1 MPa) remains within acceptable limits across most of the structure, confirming the overall stability of the dam. Localized tensile stress exceedances and the associated cracking patterns, consistent with observed damage conditions, were found to be limited in extent and not critical for the global structural stability of the dam.

Overall, the results demonstrate that the proposed methodology provides a reliable approach for addressing incomplete geomorphological data in historical hydraulic

infrastructures. By integrating GIS-based reconstruction with structural modeling, the workflow ensures consistent geometric representation and reduces uncertainties related to complex and submerged morphologies. The reconstructed model directly informs the finite element model, enhancing the robustness and physical consistency of the simulations and improving the reliability of the results.

Finally, the proposed workflow is fully reproducible using widely available tools and can be easily applied to other historical hydraulic infrastructures where incomplete topographic or bathymetric data limit conventional analysis approaches. Future developments may include the integration of dedicated in situ measurements, such as bathymetric surveys, advanced remote sensing techniques, and more detailed modelling approaches supported by uncertainty-based simulations, to further refine the reconstruction of submerged geometries and improve the accuracy of the structural model.

Acknowledgments. The authors gratefully acknowledge Eng. Pierluigi Fenu for his contribution through his Bachelor's thesis, which supported the development of this research. The Municipality of Ula Tirso is also sincerely acknowledged for its collaboration and support.

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

References

1. Lopez-Bravo, C., Lopez, J.P.: Learning from medieval Italy: urban sustainability concepts from the heritage of hydraulic infrastructures. The cases of Venice, Siena and Bologna. J. Cult. Herit. Manag. Sustain. Dev. **16**(7), 111–126 (2026)
2. Abbozzo, A., Castori, G., Mirra, M., Speranzini, E.: Preserving templar heritage: a comprehensive structural assessment of the Church of San Bevignate in Perugia (Italy). Int. J. Archit. Herit., 1–25 (2026)
3. Colace, F., Limongiello, M., Lorusso, A., Pellegrino, M., Santaniello, D., Santoriello, A.: Digital twin for cultural heritage: a computational approach to predictive conservation. Digit. Appl. Archaeol. Cult. Herit. **40**, e00519 (2026)
4. Wang, L., Mao, Y., Cheng, Y., Liu, Y.: Deep learning-based diagnosing structural behavior in dam safety monitoring system. Sensors. **21**(4), 1171 (2021)
5. Fu, X., Wang, M., Zhao, G., Xu, Y., Qi, Y., Gu, G.: A method for characterizing and analyzing the structural behavior of concrete dams in cold regions. Sci. Rep. **16**, 2660 (2026)
6. Oliveira, S., Alegre, A., Carvalho, E., Mendes, P., Proença, J.: Seismic and structural health monitoring systems for large dams: theoretical, computational and practical innovations. Bull. Earthq. Eng. **20**, 4483–4512 (2022)
7. Fedele, R., Maier, G., Miller, B.: Health assessment of concrete dams by overall inverse analyses and neural networks. Int. J. Fract. **137**, 151–172 (2006)
8. Bretas, E.M., Lemos, J.V., Lourenco, P.B.: A DEM based tool for the safety analysis of masonry gravity dams. Eng. Struct. **59**, 248–260 (2014)
9. Nguyen-Tuan, L., Könke, C., Bettzieche, V., Lahmer, T.: Numerical modeling and validation for 3D coupled-nonlinear thermo-hydro-mechanical problems in masonry dams. Comput. Struct. **178**, 143–154 (2017)
10. Robles, L., Pena, F.: Influence of the mechanical properties of masonry in the structural behavior of gravity dams. Int. J. Archit. Herit. **16**(3), 415–427 (2022)

11. Chen, J. et al.: Study on inspection and recycling evaluation of stone masonry of old Kada dam after 46 years service. Case studies. Constr. Mater. **21**, e03545 (2024)
12. Associazione Italiana di Elettrotecnica, Elettronica, Automazione, Informatica e Telecomunicazioni (AEIT): L'Elettrotecnica, ISSN 1825-828X, vol. 103. AEIT (2017)
13. Careddu, N., Grillo, S.M.: "Trachytes" from Sardinia: Geoheritage and current use. Sustainability. **11**, 3706 (2019)
14. QGIS: Development Team: QGIS Geographic Information System (version 3.34). Open Source Geospatial Foundation Project. https://www.qgis.org. Accessed 01 Jan 2026
15. Agisoft Metashape Professional (version 2.2.2), Agisoft LLC, https://www.agisoft.com. Accessed 01 Jan 2026
16. Nocerino, E., Menna, F.: Photogrammetry: linking the world across the water surface. J. Mar. Sci. Eng. **8**, 128 (2020). https://doi.org/10.3390/jmse8020128
17. Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., Reynolds, J.M.: Structurefrom-motion' photogrammetry: a low-cost, effective tool for geoscience applications. Geomorphology. **179**, 300–314 (2012)
18. Klapa, P., Żygadło, A., Pepe, M.: 3D heritage reconstruction through HBIM and MultiSource data fusion: geometric change analysis across decades. Appl. Sci. **15** (2025). https://doi.org/10.3390/app15168929
19. Quadros, N.D., Collier, P.A., Fraser, C.S.: Integration of bathymetric and topographic LiDAR: a preliminary investigation. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences (2008)
20. Yang, F., Qi, C., Su, D., Ding, S., He, Y., Ma, Y.: An airborne LiDAR bathymetric waveform decomposition method in very shallow water: a case study around Yuanzhi Island in the South China Sea. Int. J. Appl. Earth Obs. Geoinf. **109**, 102788 (2022)
21. U.S. Geological Survey Bathymetric Surveys: USGS bathymetric data and survey information. https://www.usgs.gov/faqs/where-can-i-find-bathymetric-data. Accessed 01 Jan 2026
22. NOAA National Bathymetric Source/Hydrographic Surveys: NOAA NOS hydrographic survey data. https://www.ncei.noaa.gov/products/nos-hydrographic-survey. Accessed 01 Mar 2026
23. GEBCO Shallow Water Bathymetry: GEBCO Global Bathymetric Grid and Shallow Water Data. https://www.gebco.net/data-products/gridded-bathymetry-data/shallow-water-bathymetry. Accessed 01 Jan 2026
24. Dey, S., Saksena, S., Merwade, V.: Assessing the effect of different bathymetric models on hydraulic simulation of rivers in data sparse regions. J. Hydrol. **575**, 838–851 (2019)
25. Regione Autonoma della Sardegna: Sardegna geoportale–regional spatial data infrastructure portal. https://www.sardegnageoportale.it/. Accessed 10 Mar 2026
26. Li, Z., Zhu, C., Gold, C.: Digital Terrain Modeling: Principles and Methodology. CRC Press, Boca Raton (2004)
27. Rhinoceros 8 user manual Robert McNeel & Associates (2025)
28. Midas FEA NX Analysis Reference (2025)
29. NTC: 2018 ministerial decree of 17 January 2018, official gazette of the Italian republic, No. 42, 20 February (2018)
30. Zucca, M., Reccia, E., Vecchi, E., Pintus, V., Dessì, A., Cazzani, A.: An evaluation of the structural behaviour of historic buildings under seismic action: a multidisciplinary approach using two case studies. Appl. Sci. **14**(22), 10274 (2024)

Integrated Geophysical Study at the Colonel Henry Anderson Morshead Burial Memorial, Valletta: Preliminary Results

Sebastiano D'Amico[1(✉)], Luca Piroddi[1], Emanuele Colica[1], Luciano Galone[1], and Mario Farrugia[2]

[1] Department of Geosciences, University of Malta, Msida MSD2080,, Malta
sebastiano.damico@um.edu.mt
[2] Fondazzjoni Wirt Artna – the Malta Heritage Trust, Vittoriosa BRG9038,, Malta

Abstract. This note covers the geophysical investigations carried out at the Colonel Henry Anderson Morshead burial memorial located within the Saluting Battery complex in Valletta. The main goal of the investigation is to digitally reconstruct the internal part of the memorial as well as the external monument. Furthermore, the derived 3D digital model was used to identify and locate cracks that need to be looked out. The area was also surveyed by means of a Ground Penetrating Radar (GPR) to map the subsoil of the ceremonial square at the city bastion. Here, the main documenting results are briefly presented and described, including the preliminary analysis of the photogrammetric and georadar digital datasets collected at the monument.

Keywords: Non-invasive diagnostics · Ground Penetrating Radar · 3D photogrammetry · Monumental site documentation

1 Introduction

Valletta's Saluting Battery sits in a commanding panoramic position overlooking the Grand Harbour (Fig. 1) and is an integral part of the defensive system built after the Great Siege of 1565. During that siege, the Ottomans had occupied the Mount Sciberras peninsula and installed powerful batteries there, which relentlessly bombarded Fort St. Angelo and the fortified cities of Vittoriosa and Senglea. Following the victory of the Order of St. John, construction of the new city of Valletta began in 1566, designed by military engineer Francesco Laparelli. On the site of an Ottoman battery, a powerful artillery platform was built to control the entire port.

The bastion of St. Peter and Paul, entrusted to the Order's Italian "language", became known as the Italian Post. The structure, quadrangular in shape and inclined outward for greater stability, included a large upper platform capable of housing numerous heavy cannon with an arc of fire not firing range of approximately 270°. In the 17th century, a covered loggia for the knights was added, later rotted away in 1775. In the first quarter of the 19th century, the upper section was transformed into a public garden (Upper

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 54–67, 2027.
https://doi.org/10.1007/978-3-032-30539-8_4

Barracca Gardens) to which a series of housed commemorative monuments and plaques were added.

The lower section houses the Saluting Battery itself: a rectangular platform with a continuous low parapet, lateral and central powder magazines. From the beginning, it had a ceremonial function: the firing of gun salutes in honor of visiting dignitaries, ships, religious and civil commemorations, as well as salvos to announce key naval and military victories or for state funerals. From 1875, a noonday gun started being fired from here to announce the precise moment of the passage of midday at which maritime chronometers on board ships were to be regulated. Correct time on board ship was crucial for finding the Longitude at sea. This service continued until 1923, when it was replaced by radio time signals.

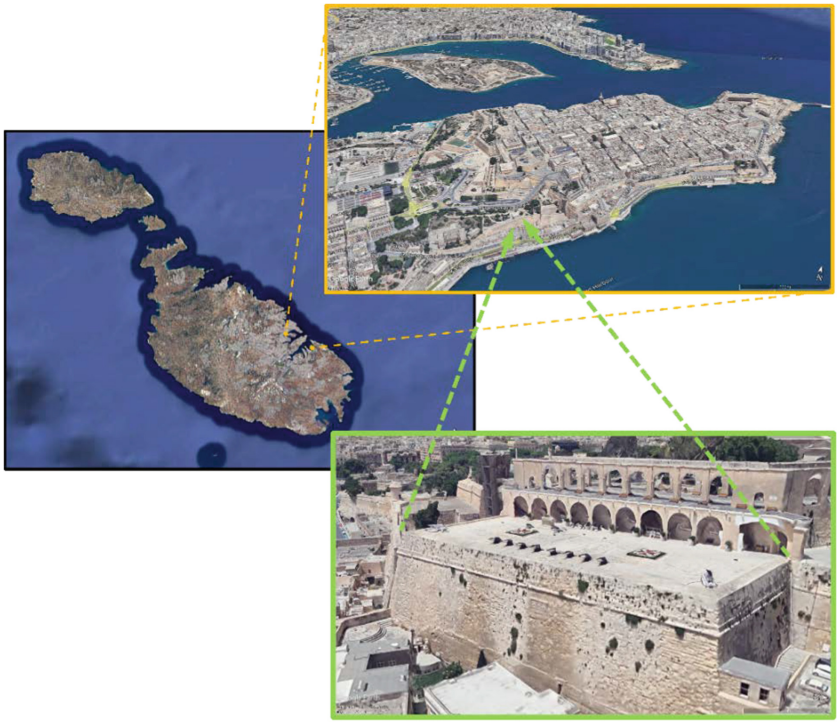


Fig. 1. Satellite view of the Maltese Archipelago, with a virtual perspective over the capital city of Valletta and, in green rectangle, of the bastion including the Upper Barracca Gardens and therein the Colonel Henry Anderson Morshead burial memorial (Google Earth source).

Over the centuries, the armament changed several times, reflecting the evolution of artillery: from the muzzle loading ordnance of the Knights' era to the French and British pieces after 1800, to the Smooth-Bore Breech Loading 32-pounders of the 19th century and the Quick Firing guns of the 20th century. During the First World War, salutes were suspended and were resumed with the Armistice of 1918.

In the 1921, following the granting of the self-government by the British agreement was reached between the civil and military authorities over the shared use of the area; the battery was partially opened to the public as part of the upper garden. During the Second World War, the guns were removed and replaced by a Bofors anti-aircraft gun; in 1942, bombings caused much damage, but Colonel Morshead's memorial although badly shaken remained standing. After a brief postwar revival, the battery was closed in 1948 and remained abandoned until 1971.

The Saluting Battery was restored in 2004 on the initiative of Fondazzjoni Wirt Artna – the Malta Heritage Trust. In 2011, eight working replicas of 32-pounder SBBL guns were installed for regular ceremonial firing. Today, the Saluting Battery is one of Valletta's main attractions: the noon-day gun and ceremonial salvos keep a centuries' old tradition alive, combining historical entertainment with the valorization of Maltese heritage.



Fig. 2. Pictures of the experimental data collection: views towards the Upper Barracca Gardens **(a)** and the Colonel Henry Anderson Morshead burial memorial **(b)**; acquisition of close-range photogrammetric data inside the memorial with a customized instrumentation **(c)** and acquisition of GPR data over the monumental square **(d)**.

This paper illustrates the geophysical investigations carried out at the Colonel Henry Anderson Morshead memorial and the main square located within the Saluting Battery complex in Valletta (Fig. 2). The main goal is to digitally reconstruct the external/internal parts of the memorial as well as the potential presence of cultural/historical features below the square pavement. Furthermore, the derived 3D digital model was used to

identify and locate cracks that need to be looked out. The balcony area was also surveyed by the mean of a Ground Penetrating Radar (GPR). The results related to the two analyses are here preliminary presented as the data interpretation is not completed as yet, and planned deeper and more detailed processing is expected to constrain general and local interpretation.

2 Methodology

2.1 Photogrammetric Survey and 3D Digital Model Reconstruction

In recent years, methodologies, technologies, approaches and software for creating digital models have undergone significant advancement and achievements. Acquisition time, post-processing tools and operational cost have also improved considerably, contributing to wider the use of such technologies in several fields. Digitization now allows reliable coverage of artefacts, ranging from small gauge (e.g. an archaeological prehistoric stone tool) to larger scales (e.g. building, urban segments or habitat areas), [1]. Photogrammetry can represent a key factor and in recent years it has been applied to planning, recording, reconstruction, and revitalization of world heritage sites [2–5]. It can also be used for supporting open-pit mine excavation stages and to plan rehabilitation strategies, as well as for forestry and management planning and disaster management [6–11]. One of the 3D digitization and reconstruction techniques is better known as photogrammetry which provides methods to obtain quantitative information. Photogrammetry can be described as the science to obtain reliable information about the spatial properties of land surfaces and objects, without physical contact [12–14]. By using reference terrain coordinates, it is possible to determine distance, area and other geographical information, as well as create detailed cartography. A further advantage of photogrammetric matching techniques is the fact that they are relatively fast and cheap in respect to the traditional laser scanner techniques [15], which are nowadays widely used in terrain models generating large amounts of 3D point data. To build a 3D model using photogrammetry, many types of images can be used such as satellite, airborne, balloon, UAVs (Unmanned Aerial Vehicles), terrestrial and even underwater images. In recent years, UAV photogrammetry imaging applications have increased rapidly due to the lower costs of a UAV for aerial surveying associated with the relative simplicity in operating such systems. With GNSS equipped drones, digital cameras and powerful computers, surveys with an accuracy of 1 to 2 cm are possible [16–20]. The 3D model of the study area has been developed by using the photogrammetry software Agisoft Metashape and the georeferencing process was carried out by using several ground control points collected through a GNSS survey for the topographic survey.

The processing of the images taken by the drone survey follow 4 main stages as sketched in Fig. 3 and summarized hereafter.

1. *Image acquisition and selection*: Images were acquired through the use of a digital camera installed on the drone assuring a good amount of image overlapping. Preliminary processing was carried out in order to discard those of poor quality.
2. *Camera alignment*: At this stage common points are selected on all the photo kept for the processing and matches through the use of photogrammetry software Agisoft

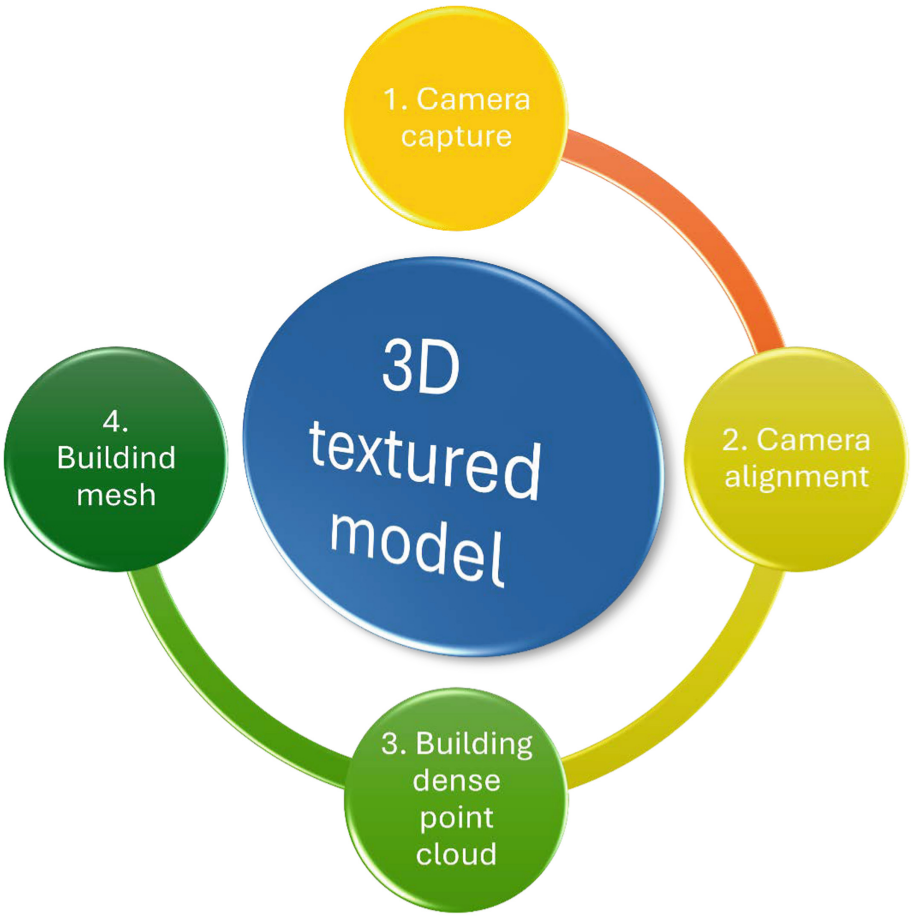


Fig. 3. Workflow processing to obtain the final 3D model.

Metashape. It also identified the position of the camera for each picture and refined camera calibration parameters. As a result, a sparse point cloud and a set of camera positions were obtained.

3. *Building dense point cloud*: The third stage involved the building of a dense point cloud based on the estimated camera positions corrected by the addition of ground control points to georeference the point cloud. If necessary, dense point cloud may be further edited and classified prior to export or proceeding to 3D mesh model generation.
4. *Building the mesh*: At this stage, the software was used to reconstruct a 3D polygonal mesh representing the object surface based on the dense or sparse point cloud according to the user's choice. Some corrections, such as mesh decimation, removal of detached components, closing of holes in the mesh, smoothing, were performed. Building texture: After the mesh geometry was reconstructed, the texture procedure

has been performed and the 3D model was created. It can be exported in several digital formats.

2.2 Ground Penetrating Radar

A wide range of geophysical techniques has been fruitfully applied to historical and cultural heritage sites, including electrical resistivity, frequency- and time-domain induction electromagnetics, magnetics, ground penetrating radar, gravity, seismics, radiometric surveys, thermography, multispectral, ground-based radar interferometry, and aerial and satellite remote sensing [21, 22].

Among them, ground-penetrating radar (GPR) is one of the most reliable, flexible and diffused methods thanks to its effectiveness (provided suitable environmental conditions) in identifying shallow natural and anthropogenic features with a high spatial resolution [3, 23–32]. GPR techniques have different on-field implementations, which mainly differ in the use of reflected or transmitted electromagnetic (EM) signals to study inaccessible domains of the medium. When only one surface of investigation is available, the most usual applications involve the use of zero- or quasi-zero-offset acquisitions by moving the GPR device in contact with the ground surface and recording the electromagnetic signals backscattered by the underlying scenario [32]. Reflections, refractions, and diffractions are generated in correspondence to the variations in the electromagnetic properties of the materials, producing the differences in velocity between the crossed subdomains [33, 34]. When a target inside the background medium has edges or dimensions comparable with the radar wavelength (that is, related to the velocity of the signal), diffraction is produced and electromagnetic waves are scattered in various directions. Even the minimum size of a recognizable target (spatial resolution) strongly depends on the wavelength of the electromagnetic signal, but it is also a function of the depth, having better resolutions for the shallowest domains [33, 34]. Signal attenuation across a medium is mostly dependent from the frequency of each component of the EM waves and to the properties of the crossed material, especially its electrical conductivity [33, 34].

GPR is de-facto a primary method for inspecting paved surfaces, in particular historical and precious pavements to search for past uses signs, archaeology, voids and possible metallic targets, like the expected scenario of the investigated semi-space at the Saluting Battery bastion. GPR can map buried foundations, ancient walls, and old road levels hidden beneath current paving. It has been used successfully in complex sites to reconstruct hidden architectural features [3, 4, 23–32, 35]. Furthermore, the GPR technology is extremely sensitive to changes in the physical properties of the subsurface, making it easy to identify air- or sediment-filled voids, like in the cases of buried rooms, crypts or burials. GPR systems detect metal objects (such as historical artifacts, pipes, or reinforcement) with high precision due to their high electromagnetic reflectivity and a particularly high Signal-to-Noise response.

The technology behind the probing system employed for the case in study is the Stepped-Frequency Continuous-Wave (SFCW) GPR, a subsurface imaging technology that transmits a sequence of discrete frequencies in linear “steps” rather than a single broadband pulse (Fig. 4). Unlike traditional Pulsed GPR, which operates in the time domain, SFCW GPR operates in the frequency domain:

1. *Sequential Stepping*: The radar sends out a sine wave at one frequency, then moves to the next, covering a wide range (Ultra-Wideband).
2. *Amplitude & Phase*: For every step, the receiver measures the amplitude and phase of the reflected signal.
3. *Synthetic Pulse*: Using an Inverse Discrete Fourier Transform (IDFT), the system converts the collected frequency data back into the time domain to create a high-resolution image of the subsurface.

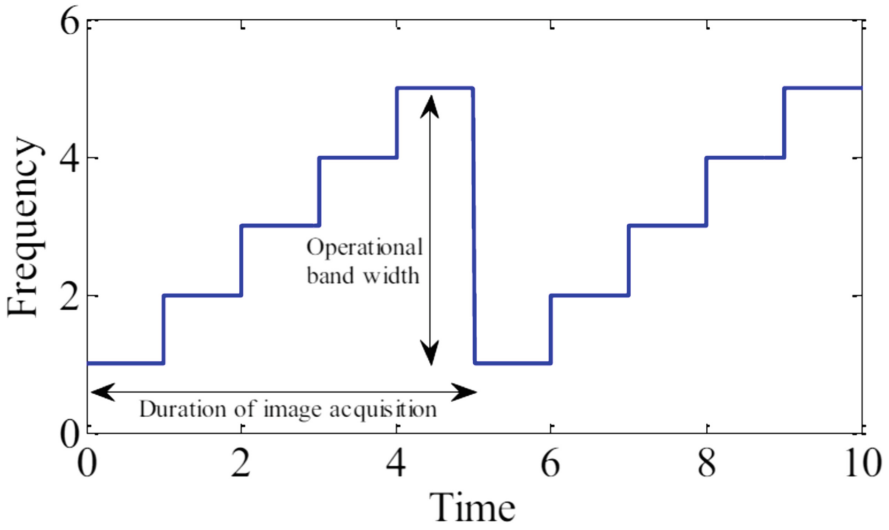


Fig. 4. Illustrative spectrum of a Stepped Frequency Continuous Wave signal (from [36]).

The experimental GPR data collection has been carried out with a zero-offset antenna setup constituted by a Screening Eagle SFCW GPR, model Proceq GS8000 Pro, with a modulated frequency over range of 40–3440 MHz, odometer and a multiband GNSS positioning system (including GPS + Glonass + Galileo + Beidou constellations). GPR system was controlled during the acquisition by the operator through a tablet connected to the internet for immediate data upload to a processing cloud environment provided by the GPR manufacturer with the availability of the most common processing steps, including band-pass filtering, background removal, gain correction, data interpolation and 3D C-scans production.

3 Data Acquisition, Processing and Results

3.1 Photogrammetry and 3D Digital Model

In this study, a small endoscopic camera and a LED light system were used in order to maximize the safety (towards the monument) during data acquisition as well as make use of the small opening realized specifically for the survey. The images were acquired

keeping the camera at the minimum focal length, while the image resolution was set at the highest level in order to acquire good quality textures. The images were taken from different-looking directions and about 80% forward overlap and 75% side overlap.

Throughout the area of interest, several markers were placed on the ground and the walls of the monument. The use of the marker was necessary to ensure correct and precise spatial-referencing of the area in order to properly export the final 3D to scale from which all the mapping and measurements have been taken. We performed four separate photogrammetry surveys and the 3D model of the study area has been developed by using the photogrammetry software Agisoft Metashape and the spatial-referencing and scaling process was carried out by using ground control points collected. More details about the processing can be found in [18]. A 3D laser distometer was used to collect the exact position of the markers and to enhance the topographical survey, which helped to produce the requested CAD (computer-aided design) drawing which will accurately detail the surveyed points. The site survey and the final CAD allows to produce a wide range of information which can be used to produce drawings. The survey results can be used for:

- Drawings – Plans, elevations, sections and lighting plans;
- 3D Point Clouds – The information can be used to produce 3D models of a site;
- Photographs – Useful for people who are looking for information about a site;
- Schedules – Including mapping of hidden structures/features

The monument was also surveyed by the use of a small Lidar instrument which is a method for determining variable distance by targeting an object or a surface with a laser and measuring the time for the reflected light to return to the receiver. It can also be used to make digital 3D representations of areas.

Figure 5 shows the 3D-derived model of the entire structure from different angles. It is possible to notice that the coupling of all the data permitted to create a unique 3D digital model which includes also the crypt underneath the “saluting battery square”. By using the 3D model we were able to map, with high precision, the location, as well as the thickness, of several features. In this case, we measured, as an example, the size of the coffin as well as of the internal part of the crypt.

3.2 Ground Penetrating Radar

The experimental GPR data collection has been carried out with a zero-offset SFCW GPR system, namely the Proceq GS8000 Pro model (40–3440 MHz). Data were acquired in a free-path fashion, following the automatic real-time generation of brief amplitude maps while continuing walking, allowing the operator to correct their path and fill the eventual empty areas between the lines distant more than the interpolation radius. After the campaign collection activity is closed and data is sent to the processing cloud environment provided by the GPR manufacturer, they are processed following the most diffuse standards, including band-pass filtering, background removal, zero-timing and gain correction, velocity analysis and time-to-depth correction, migration, Hilbert-transform, data interpolation and 3D C-scans production. In the following Fig. 6, a global view of the processed 3D dataset is provided, with color scale indicating the instantaneous amplitudes of the GPR signals, and, laterally, a histogram of the statistical distribution

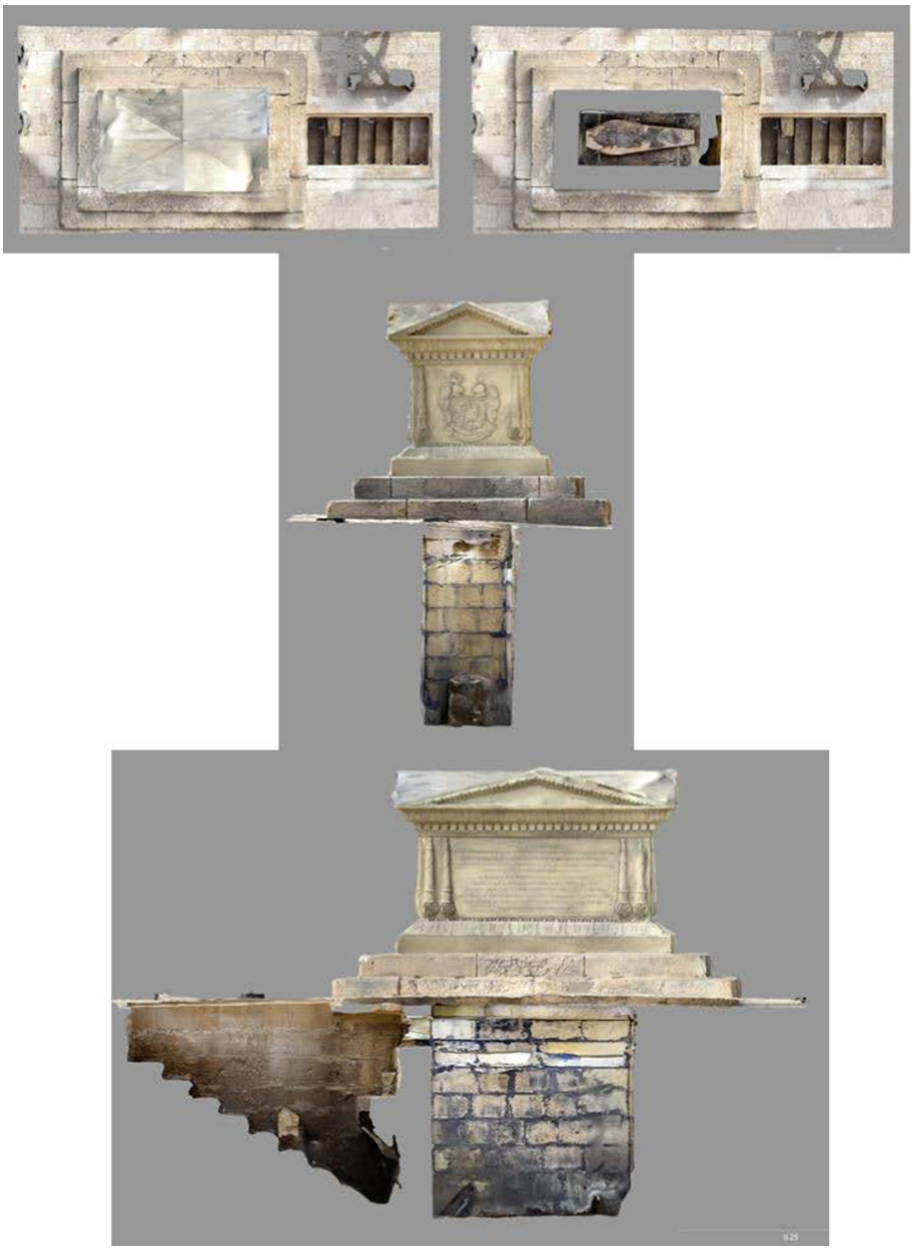


Fig. 5. Images of the 3D digital model of the memorial monument and its underneath structure.

of each amplitude value. It is possible to evaluate how much stretched the color scale is, and the quite regular distributions of the red tones indicating the few most energetic radar events.

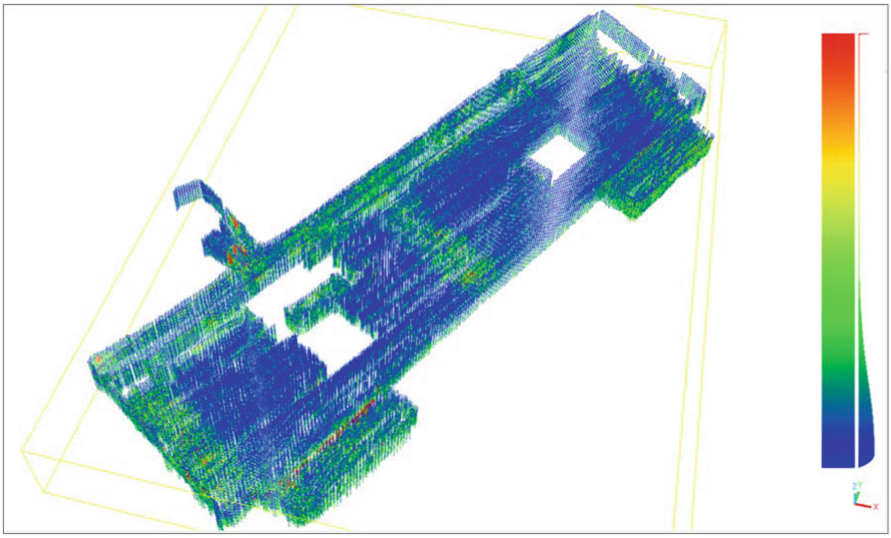


Fig. 6. Volumetric distribution of the GPR signals below the monumental square; color-bar represents the amplitudes range of the recorded wavelets after the data processing and the histogram shows the data statistical distribution along the same amplitudes range.

The graphical interpretation of preliminary GPR results is given in the following Fig. 7, where, with a different palette from the previous Fig. 6, amplitude scale is produced in a monotone ramp from white/light-gray (low amplitudes voxels) to dark reds (high amplitudes). The palette extension and coloring is a little less effective in highlighting short-range patterns but allows to clearly identify large areas/volumes with peculiar values ranges which can possibly be attributable to the signatures of different historic pavement finishes which are currently invisible at the surface due to the modern restorations (green highlighted volumes). Two large volumes correpsonding to the areas bordering the current cannoon battery also present a distinctive medium-to-high amplitude response, and can be attributable to the remaining signature of the ancient pavement or to different structural composition of the bastion at its limits. The most reflective targets noticed in the previous Fig. 6 are globally still recognizeable and locally even clearer and more contrasted with regards to the background. The orange dashed area/volume close to the colonnade seems to be the most interesting and candidate to further detailed analysis and investigations.

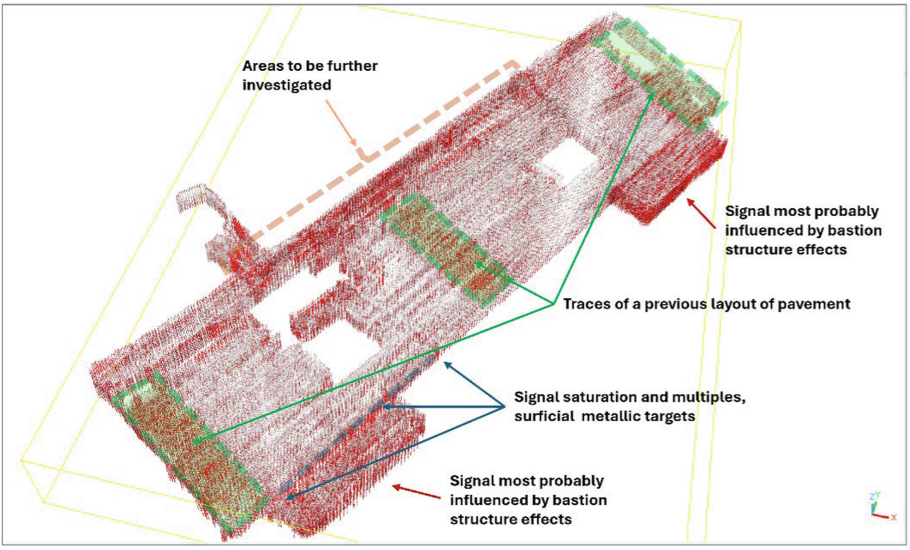


Fig. 7. Preliminary interpretation of the volumetric distribution of the GPR signals below the monumental square (color-scale from bright, low amplitudes, to dark red, high amplitudes).

4 Concluding Remarks

The application of geomatics and geophysical techniques, that is becoming a diffuse integration in high-level diagnostic investigations of Cultural Heritage sites, demonstrated here at the Colonel Henry Anderson Morshead burial memorial and the Saluting Battery square (bastion of St. Peter and Paul in Valletta, Malta) the capacity to reliably and deeply document the precious assets in detailed 3D models of visible and invisible parts through a low-invasive approach. Thanks to a customized acquisition setup, the photogrammetric survey allowed, among other results, to detect internal structural defects of the underground burial room, that we plan to map and fully interpret in a next detailed analysis. The ground penetration radar allowed to reconstruct the reflective behavior of the first volumes below a large and precious paved monumental area in a totally non invasive fashion: the preliminary interpretation of GPR datasets possibly points to the signs of the past and/or recent uses of the area (from a historic pavement remains, to cannonballs and bullets, from superficial metallic features to structural components of the bastion). The current analysis -which needs to further detailing- currently excludes the presence of large burials, but a large part (around 10% of the surveyed area) needs for additional investigation due to a diffuse -apparently incoherent- high-amplitude GPR signal response.

Acknowledgments. The authors are grateful to the Superintendence of Cultural Heritage and the Fondazzjoni Wirt Artna for supporting and permitting the study of the Colonel Henry Anderson Morshead burial memorial and Saluting battery monumental square, Valletta.

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

References

1. D'Amico, S. et al.: 3D survey and GPR for cultural heritage: the case study of SS Pietro and Paolo Church in Casalvecchio Siculo. Kermes. **107**, 11–15 (2017)
2. Al-Ruzouq, R.I., Al-Rawashdeh, S.B., Tommalieh, O.A., Ammar, M.S.: Photogrammetry and GIS for three-dimensional modeling of the dome of the rock. Appl. Geomat. **4**(4), 257–267 (2012)
3. Piroddi, L. et al.: Multi-method and multi-depth geophysical data integration for archaeological investigations: first results from the Greek City of Gela (Sicily, Italy). Remote Sens. **17**(21), 3561 (2025)
4. Ranieri, G., Trogu, A., Loddo, F., Piroddi, L., Cogoni, M.: Digital museum from integrated 3D aerial photogrammetry, laser scanner and geophysics data. In: 24th European Meeting of Environmental and Engineering Geophysics, vol. 2018, pp. 1–5. European Association of Geoscientists & Engineers (2018)
5. Piroddi, L., Calcina, S.V., Trogu, A., Vignoli, G.: Towards the definition of a low-cost toolbox for qualitative inspection of painted historical vaults by means of modified DSLR cameras, open source programs and signal processing techniques. In: International Conference on Computational Science and its Applications, pp. 971–991. Springer International Publishing, Cham (2020)
6. Esposito, G., Mastrorocco, G., Salvini, R., Oliveti, M., Starita, P.: Application of UAV photogrammetry for the multi-temporal estimation of surface extent and volumetric excavation in the Sa Pigada Bianca open-pit mine, Sardinia, Italy. Environ. Earth Sci. **76**(3), 103 (2017)
7. Mafanya, M., Tsele, P., Botai, J., Manyama, P., Swart, B., Monate, T.: Evaluating pixel and object based image classification techniques for mapping plant invasions from UAV derived aerial imagery: Harrisia pomanensis as a case study. ISPRS J. Photogramm. Remote Sens. **129**, 1–11 (2017)
8. Vetrivel, A., Duarte, D., Nex, F., Gerke, M., Kerle, N., Vosselman, G.: Potential of multi-temporal oblique airborne imagery for structural damage assessment. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. **3**, 355–362 (2016)
9. Leucci, G. et al.: Stability assessment and geomorphological evolution of sea natural arches by geophysical measurement: the case study of Wied il-Mielah window Gozo, Malta. Sustainability. **13**(22), 12538 (2021)
10. Galone, L. et al.: Evolution of coastal cliffs characterized by lateral spreading in the Maltese archipelago. Remote Sens. **16**(16), 3072 (2024)
11. Colica, E. et al.: Rockfall hazard assessment using UAV photogrammetry and 3D modelling at Selmun cliff in Malta. Discov. Geosci. **3**(1), 240 (2025)
12. Gomez, C., Purdie, H.: UAV-based photogrammetry and geocomputing for hazards and disaster risk monitoring–a review. Geoenviron. Disasters. **3**(1), 23 (2016)
13. Luhmann, T., Robson, S., Kyle, S., Boehm, J.: Close-Range Photogrammetry and 3D Imaging. de Gruyter, Berlin (2014)
14. Durou, J.D., Falcone, M., Quéau, Y., Tozza, S. (eds.): Advances in Photometric 3d-Reconstruction, vol. 3. Springer (2020)
15. Schenk, T.: Introduction to Photogrammetry. Department of Civil and Environmental Engineering and Geodetic Science The Ohio State University, Columbus, OH, USA (2005)
16. Nuikka, M. et al.: Comparison of three accurate 3D measurement methods for evaluating as-built floor flatness. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 37, pp. 129–134, Beijing (2008)
17. Corrigan, F: Introduction to UAV Photogrammetry and Lidar Mapping Basics (2017). https://www.dronezon.com/learn-about - %20drones - quadcopters / introduction - to -uav-photogrammetry-and-lidar-mapping-basics

18. Colica, E., Micallef, A., D'Amico, S., Cassar, L.F., Galdies, C.: Investigating the use of UAV systems for photogrammetric applications: a case study of Ramla Bay (Gozo, Malta). J. Malta Chamber Sci. **5**, 125–131 (2017)
19. Pfeiffer, R. et al.: Use of UAVs and deep learning for beach litter monitoring. Electronics. **12**(1), 198 (2022)
20. Colica, E. et al.: Using unmanned aerial vehicle photogrammetry for digital geological surveys: case study of Selmun promontory, northern of Malta. Environ. Earth Sci. **80**(17), 551 (2021)
21. Piroddi, L. et al.: Imaging cultural heritage at different scales: part I, the micro-scale manufacts. Remote Sens. **15**, 2586 (2023)
22. Piroddi, L. et al.: Imaging cultural heritage at different scales: part II, the Meso-scale sites. Remote Sens. **17**, 598 (2025)
23. Papadopoulos, N., Sarris, A., Yi, M.J., Kim, J.H.: Urban archaeological investigations using surface 3D ground penetrating radar and electrical resistivity tomography methods. Explor. Geophys. **40**, 56–68 (2009) [Google Scholar] [CrossRef]
24. Conyers, L.B., Leckebusch, J.: Geophysical archaeology research agendas for the future: some ground-penetrating radar examples. Archaeol. Prospect. **17**, 117–123 (2010) [Google Scholar] [CrossRef]
25. Piro, S., Campana, S.: GPR investigation in different archaeological sites in Tuscany (Italy). Anal. Comp. Obtained results. Near Surf. Geophys. **10**, 47–56 (2012) [Google Scholar] [CrossRef]
26. Zhao, W., Forte, E., Levi, S.T., Pipan, M., Tian, G.: Improved high-resolution GPR imaging and characterization of prehistoric archaeological features by means of attribute analysis. J. Archaeol. Sci. **54**, 77–85 (2015) [Google Scholar] [CrossRef]
27. Piroddi, L., Calcina, S.V., Fiorino, D.R., Grillo, S., Trogu, A., Vignoli, G.: Geophysical and remote sensing techniques for evaluating historical stratigraphy and assessing the conservation status of defensive structures heritage: preliminary results from the military buildings at San Filippo bastion, Cagliari, Italy. In: International Conference on Computational Science and its Applications, pp. 944–959. Springer International Publishing, Cham (2020, July)
28. Trinks, I. et al.: Large-area high-resolution ground-penetrating radar measurements for archaeological prospection. Archaeol. Prospect. **25**, 171–195 (2018) [Google Scholar] [CrossRef]
29. Persico, R. et al.: GPR investigations at St John's co-Cathedral in Valletta. Near Surf. Geophys. **17**, 213–229 (2019) [Google Scholar] [CrossRef]
30. Colica, E. et al.: GPR Investigation at the Archaeological Site of Le Cesine, vol. 12, p. 412. Information, Lecce, Italy (2021) [Google Scholar] [CrossRef]
31. Manataki, M., Vafidis, A., Sarris, A.: GPR data interpretation approaches in archaeological prospection. Appl. Sci. **11**, 7531 (2021) [Google Scholar] [CrossRef]
32. Piroddi, L., Rassu, M.: Application of GPR prospection to unveil historical stratification inside monumental buildings: the case of San Leonardo de Siete Fuentes in Santu Lussurgiu, Sardinia, Italy. Land. **12**, 590 (2023) [Google Scholar] [CrossRef]
33. Jol, H.M.: Ground Penetrating Radar Theory and Applications, 1st edn. Elsevier, Amsterdam (2008) [Google Scholar]
34. Reynolds, J.M.: An Introduction to Applied and Environmental Geophysics. Wiley, Hoboken, NJ (2011)
35. Piroddi, L., Vignoli, G., Trogu, A., Deidda, G.P.: Non-destructive diagnostics of architectonic elements in San Giuseppe Calasanzio's church in Cagliari: a test-case for micro-geophysical methods within the framework of holistic/integrated protocols for artefact knowledge. In: 2018 Metrology for Archaeology and Cultural Heritage MetroArchaeo, pp. 17–21. IEEE (2018)

36. Piroddi, L., Calcina, S.V.: Integrated vibration analysis for historical dome structures: a complementary approach based on conventional geophysical methods and remote sensing techniques. In: Gervasi, O. (ed.) Computational Science and Its Applications—ICCSA 2020 Lecture Notes in Computer Science, vol. 12255. Springer, Cham (2020)

Investigating the Castello Archaeological Park (Cersosimo, Italy) Using a Combination of Geophysical Prospections, Aerial and Terrestrial Photogrammetry

Marilena Cozzolino[1(✉)], Vincenzo Gentile[1], and Paolo Mauriello[2]

[1] Department of Biosciences and Territory, University of Molise, C.da Fonte Lappone, 86090 Pesche, Italy
marilena.cozzolino@unimol.it

[2] Department of Agricultural, Environmental and Food Sciences, University of Molise, 86100 Campobasso, Italy

Abstract. The Castello Archeological Park is located in the Municipality of Cersosimo in the Basilicata Region (Province of Potenza, southern Italy). Significant questions remain unanswered despite earlier archaeological studies that have clarified a number of important aspects of the settlement, including the layout of the southeastern sector of the enclosure wall, the location and the spatial boundaries of two major architectural complexes, and the identification of the southern access gate to the fortified area. Specifically, there are still interrogations about the settlement's internal spatial layout and the complete scope of the defensive circuit. In this article, the initial findings of a multidisciplinary study program aimed at increasing the site's current understanding by combining high-resolution photogrammetric survey and non-invasive geophysical prospection are presented. The study intends to improve knowledge of the settlement's structural layout, find previously overlooked architectural features, and create a thorough geospatial framework to direct further archaeological research by fusing subsurface diagnostic methods with three-dimensional surface documentation.

Keywords: Castello Archaeological Park · Cersosimo · Aerial and terrestrial photogrammetry · Geophysical prospections

1 Introduction

At an elevation of about 723 m above sea level, the Castello Archaeological Park of is located in the upper sector of the Municipality of Cersosimo inside the Province of Potenza (Basilicata, southern Italy) (40°02′42.27′N; 16°21′06.05′E) (Fig. 1). The site is strategically placed on a hill with a view of the Saramento valley, the Lappio stream, and the natural route that connects the inland regions to the Ionian coastal plain and passes through the current center of Oriolo [1]. In line with larger settlement dynamics recorded in Basilicata and northern Calabria during the pre-Roman era, this topographical configuration amply demonstrates the settlement's defensive and territorial control

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 68–85, 2027.
https://doi.org/10.1007/978-3-032-30539-8_5

function. The settlement's location guaranteed visual command over communication corridors and hydrographic networks.

A fortification system that takes advantage of the terrain's natural shape defines the archeological site. The defensive circuit forms an irregular quadrilateral with a total circumference of around 900 m by adjusting to contour lines and slope breaks. Significant elements of the fortification are still recognizable, despite the fact that some of it has been destroyed by artificial disturbances, vegetation growth, and natural erosion. The southwest corner of the enclosure, where excavations conducted by Lacava in 1969 revealed a postierla, a secondary access gate to the ancient city, corresponds to the best-preserved sector [2, 3] (Fig. 2a, b). Because it sheds light on the settlement's defense tactics and circulation patterns, this feature is very interesting from an architectural standpoint. Furthermore, the first systematic documentation and archaeological investigations of the polygonal masonry walls of Castello Hill was achieved describing the building methods and spatial arrangement of the fortification system. In order to establish the spatial extent of the archaeological remains and confirm the structural continuity of the defensive circuit, De La Genière, on behalf of the Superintendence of Antiquities of Basilicata, conducted further research in 1969 [4]. These excavations greatly expanded our understanding of the settlement's layout by establishing the fortress walls' adaptability to the local geomorphology and recording important architectural features.

Two major architectural complexes, commonly known as Building A and Building B, have been found within the walled perimeter [5]. They are both thought to date to the third century B.C. One of the best examples of a residential complex is Building A, which has a central peristyle courtyard surrounded by symmetrically placed rooms (Fig. 2c, d) [6]. A peristyle court is a comparatively advanced architectural typology that reflects both socioeconomic level and functional structure suggesting a cohesive urban project. Building B was likewise built on a rectangular ground plan. Its architectural arrangement implies either a supplemental residential role or a structure with administrative or community objectives within the walled village, despite the fact that it is less well recorded.

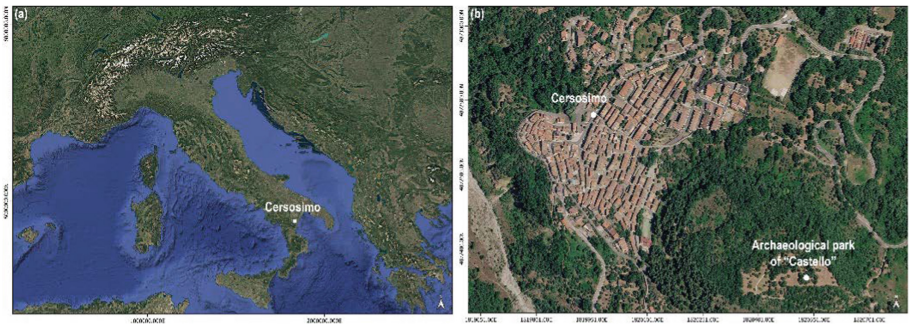


Fig. 1. Location of Cersosimo (PZ) on a Google Earth™ satellite image of Italy (a), location of the Castello Archaeological Park (b).

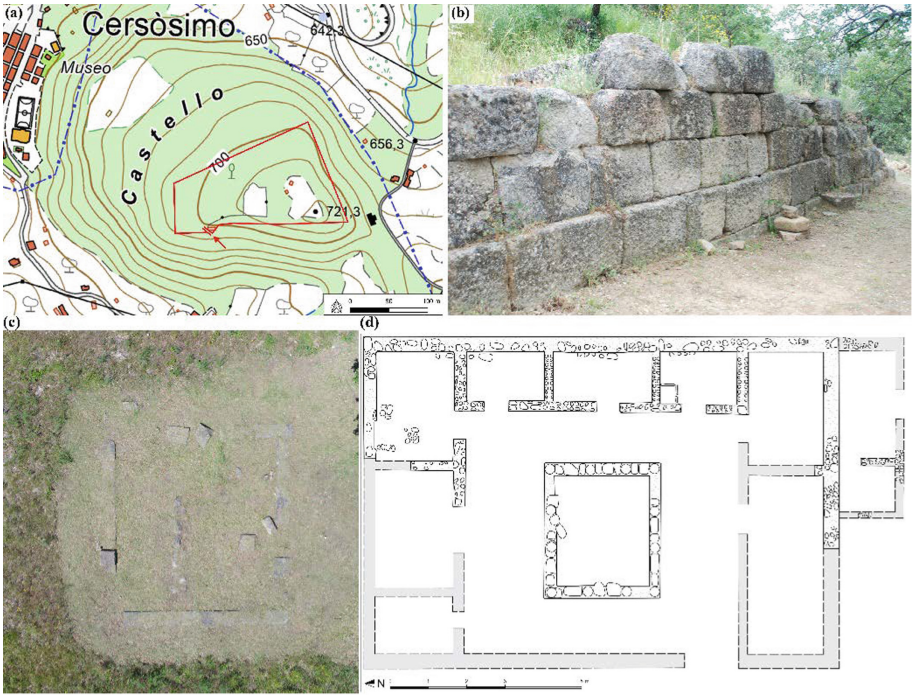


Fig. 2. Wall circuit (a), pòstierla (secondary access gate to the ancient city) (b), peristyle of Building A (c), vector survey and reconstructive hypothesis of Building A (d).

At least three main stages of habitation have been recognized by archaeological research. The first two stages are ascribed to the third century B.C., which coincided with the maximum stage of architectural investment and urban expansion. The site seems to have been abandoned after this prosperous period, most likely as a result of larger geopolitical changes brought about by Roman conquest in southern Italy. The consolidation of Roman rule resulted in the reorganization of territorial control structures and the progressive decline of pre-existing hilltop fortifications, as seen in many fortified communities throughout Basilicata and Calabria. Only in the Middle Ages Castello was interested by a return to habitation, albeit at a far lower intensity than during the previous stages.

Even with the significant contributions of earlier archaeological expeditions, there are still some features of the site that are not fully understood. There are still questions about the complete extent of the defensive circuit, even though the southeast section of the enclosure wall has been fairly thoroughly documented, including the location of the southern entry gate and the boundaries of Building A and Building B. Specifically, the wall's continuation toward the northwest and north has not been proven. The identification of structural alignments in these places has been complicated by erosion processes, abundant vegetation cover and a lack of new excavation campaigns.

The interior structure of the settlement is a second unsettled matter. The spatial arrangement of other residential, commercial or community constructions is still mostly

speculative, despite the fact that the two primary buildings have been documented. More research is needed to determine whether terracing systems, open spaces, circulation routes and potential infrastructure components (such drainage systems or storage facilities) are present. Rebuilding Castello's socio-political structure, functional zoning, and settlement hierarchy within the larger regional context of fortified centers requires an understanding of the internal urban plan.

Given these unanswered problems, combining high-resolution photogrammetric data with non-invasive geophysical surveys has been proposed as methods for filling in the gaps in the archaeological record. Without immediately turning to large-scale excavation campaigns, it was possible to refine theories about the continuity of the defensive circuit and to suggest more thorough reconstructions of the settlement's internal spatial organization.

2 Integrated Geophysical and Photogrammetric Methodology

Adopting integrated and non-invasive techniques that may combine high-resolution surface recording with subsurface excavation is becoming very useful for the study of complex archaeological environments. Developments in digital photogrammetry and geophysical prospection have significantly changed the practice of archaeology by allowing researchers to identify buried structures, reconstruct spatial organization and record architectural features with a degree of accuracy and thoroughness that was previously impossible without extensive excavation [7]. These technologies support minimally damaging research methods that are in line with modern cultural preservation and sustainable site management principles, in addition to improving interpretive capacity [8].

The current study uses a multidisciplinary approach that combines aerial and terrestrial photogrammetric surveying with frequency-domain electromagnetic (FDEM) prospection. The main goal is to create an extensive spatial dataset that can help with long-term documentation and conservation planning, enhance archaeological interpretation, and direct future excavation strategies. By combining subsurface geophysical diagnostics with three-dimensional modeling and georeferenced cartographic products, the research establishes a multi-scale analytical approach that bridges landscape-level assessment and detailed architectural reconstruction.

In this context, geophysical surveys offer spatially continuous data on subsurface electrical variability, allowing to identify anomalous zones that might be connected to buried archaeological objects. Structure from Motion (SfM) and Multi-View Stereo (MVS) techniques are used in photogrammetry to produce high-resolution three-dimensional models, orthophotos, and Digital Elevation Models (DEMs) that accurately depict topography and building remnants. Consistency, metric reliability and interoperability between many information sources are ensured by integrating these datasets into a single Geographical Information System (GIS).

The theoretical background, data collection procedures, processing workflows, and analytical techniques used in this study are explained in the following sections. In order to ensure scientific reproducibility and provide a reference framework for similar applications in archaeologically complex situations, special attention is paid to methodological choices, instrumental parameters and quality control procedures.

2.1 Geophysical Survey: Inductive Electromagnetic Prospection

Before systematic excavations are realized, geophysical prospection techniques provide a non-invasive methodological framework for an indirect detection and a first characterization of archaeological deposits. Over the past few decades, different experimental applications proved the effectiveness of geophysical surveys in identifying buried anthropogenic remains from a broad range construction typologies of various chronological phases [8–11].

Magnetic methods, among the most commonly used techniques, have proven especially successful in identifying features linked to thermally altered materials, combustion structures, filled ditches, pits, and masonry remains [12–14]. This is due to the quantifiable difference in magnetic susceptibility and remanent magnetization between archaeological targets and surrounding sediments. Frequency-domain electromagnetic (FDEM) techniques have also become more and more important for the identification of subsurface heterogeneities associated with human disturbances, given their ability to quickly evaluate spatial variations in apparent electrical conductivity and magnetic susceptibility [15, 16]. Also, electrical techniques including resistivity and induced polarization surveys have been widely used to identify buried stone foundations, voids, compacted surfaces and stratigraphic discontinuities [17–19]. Furthermore, ground penetrating radar (GPR) has significantly advanced high-resolution subsurface imaging by using dielectric permittivity contrasts to recreate 2D and 3D stratigraphic geometries [20–22].

The choice of the best geophysical method depends intrinsically on a number of interconnected factors [23]. These include the geological, pedological and hydrological setting, the surface accessibility, the vegetation cover, the logistical conditions, the spatial extension of the survey area, the expected geometry, the size, the burial depth and the physical characteristics of the target features. Signal attenuation, penetration depth and anomaly resolution can all be influenced by soil texture, moisture content and natural heterogeneity. As a result, the methodology selection requires a careful initial analysis of the expected physical differences between archaeological structures and the host sedimentary matrix, in addition to an assessment of operational limitations and the intended spatial resolution.

In the examined case study, the frequency-domain electromagnetic (FDEM) method was considered to be the best approach to reach the research goals in the frame of the specific environmental circumstances. The site is characterized by superficial lithological heterogeneity consisting in an alternation of sandy and clay-rich strata. Because the clay mineral abundance and the grain-size distribution have a significant impact on electrical conductivity, these contrasts are very important for electromagnetic research. Additionally, as recorded ancient objects in the site are mostly bricks, ceramic fragments and limestone blocks, the contrasts from the nearby strata is significant and offer an advantageous framework for the use of FDEM measurements.

The exploratory aspect of the study was another point in favor of the FDEM approach. The research area was completely unexplored and had not previously undergone thorough subsurface geophysical mapping. A technique that could quickly cover large areas while retaining sufficient sensitivity to changes in macro scale conductivity was considered ideal in this situation. In this case, the main goal was to identify broad zones with notable electromagnetic abnormalities rather than to reconstruct architectural geometries in detail. Following their delineation, these macro areas could be examined using targeted excavation techniques and higher resolution methods (such as dense-grid resistivity or GPR surveys).

The FDEM approach is theoretically founded on the concepts of electromagnetic induction [24]. A primary low-frequency alternating electromagnetic field is produced by a transmitting coil and travels beneath the surface. This primary field creates eddy currents in conductive bodies when it comes into contact with them. A spatially separated receiving coil detects the secondary electromagnetic field created by these currents. The two orthogonal components of the observed secondary field are the quadrature (imaginary) component and the in-phase (real) component. While the quadrature component is mainly associated with the bulk electrical conductivity of the subsurface materials and is commonly expressed in parts per million (ppm) relative to the primary field, the in-phase response is especially sensitive to highly conductive or metallic objects and to changes in magnetic susceptibility [24].

The quadrature component and the apparent electrical conductivity of the studied volume can be linearly correlated under low-induction-number conditions. Given the hypothesized inverse relationship between conductivity and resistivity, apparent resistivity values can be obtained from this quantity. While resistivity is reported in Ohm $\times$ meters (Ω m), electrical conductivity is generally expressed in milli Siemens per meter (mS / m), which is equivalent to 10^{-3} Siemens per meter. An approximation of the two parameters' relationship is as mS / m = 1000 / Ohm m. This conversion improves the dataset's interpretability within a larger geophysical context and makes comparison with electrical resistivity survey results easier.

The GSSI Profiler EMP-400 [25], a multi-frequency FDEM instrument intended for near-surface research, was used to collect data for this study. In order to maximize sensitivity to fluctuations in shallow conductivity, measurements were taken in continuous acquisition mode with vertically aligned dipoles at various operating frequencies between 2 and 15 kHz. Given that skin depth in FDEM surveys is strongly dependent on local ground conductivity and subsurface heterogeneity, any depth-slice representation would necessarily be highly approximate and potentially misleading if presented as a quantitative depth model. For this reason, we have chosen to retain the unique 15 KHz frequency-based interpretation without introducing additional derived depth-slice products.

In order to provide uniform spatial sampling and sufficient lateral resolution for anomaly detection, five research areas were methodically covered by parallel survey profiles placed in a regular grid with 1 m spacing (Fig. 3). Previous excavation findings in the studied area show the existence of relatively small to medium-sized subterranean structures, such as pits, wall foundations, and linear features, which are usually between a few meters and a sub-meter in width. A balance between survey resolution, site coverage,

and the anticipated sizes of the archaeological features based on earlier studies led to the selection of a 1 m profile spacing.

For eliminating experimental drift and artificial noise, raw conductivity data underwent typical filtering and quality control processes during the processing phase. A despiking process based on a moving-window method was used to eliminate outliers. In particular, data points that deviated more than ±2.5 standard deviations from the local mean were classified as noise and removed from additional examination. This cut-off was chosen to balance the preservation of true subsurface variability with efficient noise reduction. After making some initial adjustments, the filtering was applied, and the outcomes were visually examined to make sure that important geophysical features had not been unintentionally eliminated. In order to make geological and archaeological interpretation easier, the apparent conductivity measurements were transformed into resistivity values. Then, the processed data were interpolated using a contouring software and were shown as two-dimensional resistivity maps. A common logarithmic scale was used for map presentation due to the wide range of resistivity values found throughout the site, which reflects significant lithological heterogeneity. This method allowed to more clearly distinguish between conductive clay rich zones and more resistive sandy or anthropogenic characteristics, as well as to better represent both low and high magnitude anomalies.



Fig. 3. FDEM acquisition areas (magenta color), postierla (gate) (A), Bulding A (B), polygonal wall (C), Building B (D), terrace wall (E).

2.2 Photogrammetric Survey

Nowadays, photogrammetry is a crucial methodological element in heritage management and archaeological site recording thanks to developments in digital imaging, computational power and automated image processing techniques [26–33]. In order to extract three-dimensional spatial information from two-dimensional image datasets, photogrammetry is fundamentally based on mathematical and computer vision techniques. Specialized algorithms reconstruct dense point clouds, estimate intrinsic and extrinsic camera parameters, and calculate the spatial coordinates of detected features within a single reference system by identifying homologous points across several overlapping images. Structure from Motion (SfM) and multi-view stereo (MVS) approaches are the main theoretical foundation of contemporary digital photogrammetry [34–37]. In order to simultaneously estimate camera locations, orientations and 3D point coordinates, SfM algorithms automatically identify unique key points in individual photos, match these features across image pairings, and repeatedly solve the bundle adjustment issue. This method improves operational flexibility in field situations by enabling the reconstruction of complicated geometry even in the absence of pre-calibrated camera systems.

Multi-scale documenting of the archaeological site was made possible by the integration of aerial and terrestrial photogrammetry inside a single SfM based approach, which combined high resolution architectural detail with landscape level spatial coverage. In order to ensure thorough multi scale recording of the research region, two separate datasets were created: 151 frames from the aerial survey and 147 frames from the terrestrial survey. In addition to supporting stratigraphic study and archaeological interpretation, this thorough digital reconstruction supports long-term heritage preservation, virtual dissemination, and quantitative spatial assessment within an interdisciplinary research framework.

2.2.1 Aerial Photogrammetry

A remotely piloted aircraft system (RPAS), namely a Yuneec H520 platform, was used to conduct aerial surveys. It was equipped with a 1-in. CMOS sensor camera with a fixed focal length of 26 mm and a resolution of 20 megapixels. To provide adaptive navigation in response to environmental constraints, that is the presence of dense forest cover and varied topographic conditions, manual flight control was implemented. A rectangular grid flying pattern was used for image acquisition in order to optimize redundancy and coverage. To enhance the quality of three-dimensional reconstruction, both nadiral and oblique images were gathered, particularly along vertical structures and morphological discontinuities. Geometric precision was maintained while operational safety was ensured by dynamically adjusting the flight altitude to avoid obstructions and maintain the ideal ground sampling distance (GSD).

2.2.2 Terrestrial Photogrammetry

A mirrorless Sony Alpha 5100 digital camera with a 23.5 × 15.6 mm CMOS sensor, a 24-megapixel resolution and a variable focal length lens (16–50 mm), was used for complementary ground-based photogrammetry. The capture of terrestrial images was primarily intended to document the old city's entrance gate, which is distinguished by vertical architectural details that necessitate meticulous reconstruction. In both longitudinal and transverse orientations, images were taken along horizontal bands with an average overlap of about 80%. This high level of redundancy is crucial to guarantee strong key point matching, reduce reconstruction gaps and enhance geometric stability during bundle modification.

2.2.3 Photogrammetric Processing Workflow

Agisoft Metashape software (LLC, St. Petersburg, Russia) [38] was used for image processing, utilizing a SfM approach combined with dense stereo matching techniques. In order to guarantee geometric accuracy, model optimization and geospatial consistency, the processing pipeline adhered to a specified set of procedures:

1. Image Alignment (Sparse Point Cloud Generation). In this first stage, the software estimated the relative camera locations and orientations and automatically identified homologous features across overlapping images. Both exterior and interior orientation parameters were determined using bundle adjustment optimization. Focal length, main point coordinates and lens distortion coefficients (radial and tangential) are examples of interior parameters that characterize the camera system's inherent optical properties. The camera's spatial position and orientation within the reference system are defined by the exterior orientation parameters, which are composed of three translational components (X, Y, and Z) and three rotational parameters (given as Euler angles). This step results in a sparse (discrete) point cloud that depicts the scanned scene's initial geometric structure (Fig. 4a).
2. Dense Point Cloud Construction. Using multi-view stereo matching, which is based on creating depth maps for overlapping picture pairs, dense reconstruction was accomplished. During bundle adjustment, depth maps were calculated using the previously obtained orientation parameters, guaranteeing geometric coherence through iterative refinement. From its associated depth map, each image contributed to partial dense point clouds. Additional noise filtering was performed to overlapping regions once these partial reconstructions were combined into a single dense point cloud. In order to maintain radiometric fidelity, color information was directly sampled from the original images, while surface normals were computed via local plane fitting inside pixel neighbors. A high-resolution 3D depiction of the surrounding landscape and archaeological buildings was produced by the dense cloud that resulted (Fig. 4b).
3. Ground Control Point (GCP) Integration and Georeferencing. GCP markers were used to ensure accurate model scaling and spatial location. A total station survey's coordinates were imported into the photogrammetric project and manually linked tothe locations of the related images. The model could be converted from a relative coordinate framework to an absolute georeferenced system by using GCPs. This stage

is crucial since model scaling, orientation, and overall spatial accuracy can be greatly impacted by mistakes in GCP identification or coordinate input (Fig. 4c).

4. 3D Mesh Generation and Optimization. Surface reconstruction methods were used to transform the dense point cloud into a triangulated surface model (mesh). Metric analysis is made easier and continuous surface representation is made possible by this modification. Optimization techniques, such as mesh smoothing and decimation filtering, were used to increase computing efficiency and decrease geometric noise. These procedures produced a lighter, easier-to-manage 3D model that is appropriate for viewing and analysis without sacrificing structural integrity by reducing the number of polygons while maintaining crucial morphological characteristics (Fig. 4d).
5. Texture Application to the Mesh. Applying color content to the mesh was the last stage. A photorealistic 3D model can be created by applying photographic photos to the model thanks to the texture algorithm. Ultimately, texture maps (Fig. 5a), a georeferenced Digital Elevation Model (DEM) (Fig. 5b, c) and orthophotos (Fig. 5d) were produced.

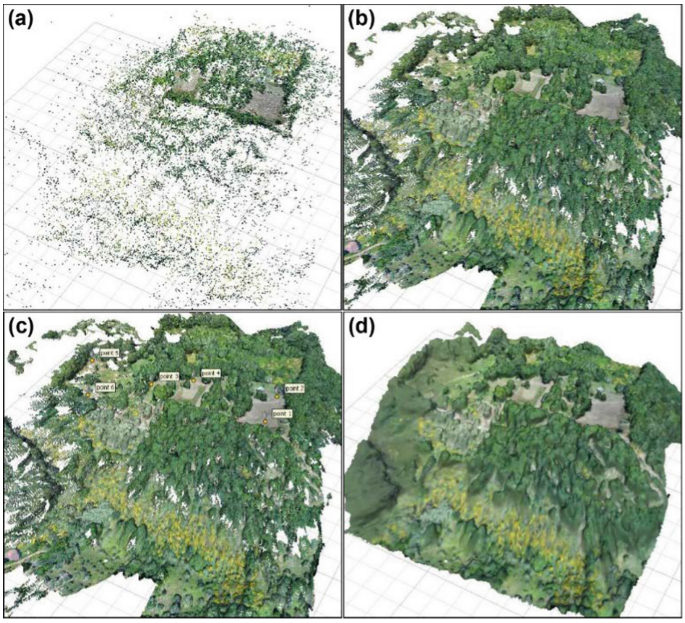


Fig. 4. Photogrammetric Processing Workflow: Sparse Point Cloud Generation (a), Dense Point Cloud Construction (b), Georeferencing Ground control points (c), 3D Mesh Generation (d).

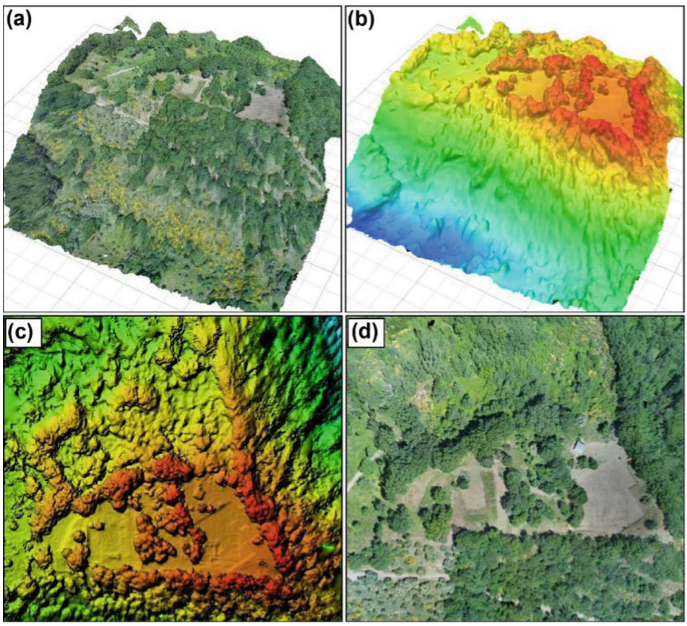


Fig. 5. Photogrammetric Processing Workflow: texture (a), digital elevation models (b, c) and ortophoto (d).

3 Results

The existing state of information regarding the archaeological site in the Castello area has been significantly improved and updated through a multi-scale analytical method that integrates data from non-invasive geophysical investigations and aerial photogrammetry obtained using remotely piloted aircraft systems (RPAS) within a single spatial framework. To ensure long-term accessibility and interoperability, all processed datasets have been georeferenced and stored within a Geographical Information System (GIS) (QGIS Development Team (2024), Open Source Geospatial Foundation Project, http://qgis.osgeo.org, accessed on 15 December 2025). Several types of spatial maps have been created and arranged as theme layers within this digital infrastructure (Fig. 6) offering a solid foundation for upcoming comparative studies and conservation condition monitoring.

To examine the spatial distribution of heights throughout the research area, a georeferenced Digital Elevation Model (DEM) was created. The detection of topographic anomalies that may be connected to defensive buildings, terrace systems and buried structure is made possible by this extensive elevation dataset. Cadastral maps, orthophotographic mosaics, and DEM-derived contour models were among the topographic bases used to create and reconcile updated archaeological plans of previously recorded remains. By combining the analysis of geophysical datasets with a comparison with established architectural patterns, interpretive designs of buried ancient sites were developed. The

resistivity distributions and electromagnetic contrasts found during the geophysical survey serve as the main foundation for these reconstructions, which are supported by morphological alignments seen in the DEM and orthophotos.

Specifically, the combined analysis reveals a number of noteworthy geographical patterns (Fig. 7). The pattern of high vegetation in the settlement's southern portion seems to follow the enclosure wall's assumed path. In archaeological settings, vegetation growth frequently serves as an indirect indicator of buried masonry by reflecting structural discontinuities and variations in subsurface moisture retention. The access gate and the previously identified excavated portion (shown in blue) are connected by the first dashed white line (Fig. 7), which suggests that the defensive structure may continue along this alignment.

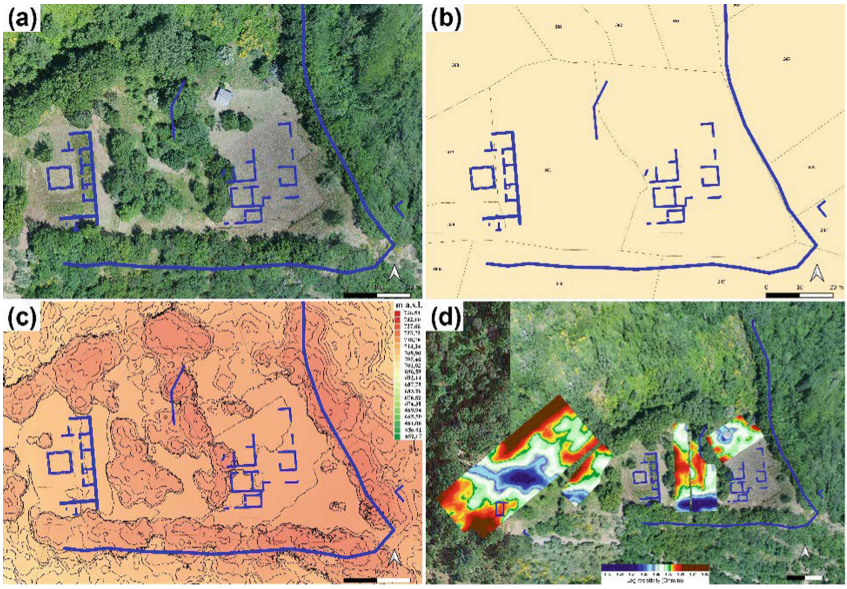


Fig. 6. Examples of spatial maps created and arranged as theme layers within Geographical Information System: orthophotographic mosaic (a), cadastral map (b), DEM-derived contour model (c) and resistivity map relative to the frequency of 15 kHz (d) with indication of previously recorded remains in blue.

The readability of traces was helped by the creation of high-resolution orthophotos of the gate (Fig. 8), the most important architectural features of the archaeological site, acquired during the terrestrial survey. The resulting outputs offer geometrically corrected and metrically precise representations of the entry gate for direct measurement, material investigation and stratigraphic reading of masonry units. These products are very useful for architectural documentation and conservation planning. The three-dimensional model was used to derive detailed architectural elevations in addition to orthophotos. The structure's true scale and geometric proportions are preserved in these elevations, which are vertically projected views of the building. Improved architectural interpretation, long-term heritage management, structural monitoring and virtual dissemination

initiatives are all aided by the incorporation of these data into the Castello area's larger GIS database.

A number of probable terrace walls, indicated by dashed white lines numbered 2 through 6, can also be found on the general interpretative map (Fig. 7). These traits are in agreement with the slope stabilization and artificial leveling methods used in fortified hilltop towns. Particularly interesting is the space between sections 4, 5 and 6. Together with the eastern border wall, these alignments produce a broad and level platform. The morphological consistency and the spatial organization of this area indicate a probable functional planning and a deliberate terracing. If the substantial scrub vegetation that currently obstructs surface vision is cleared, this area, given its size and topographic regularity, likely offers a large deal of archeological potential and merits targeted research.

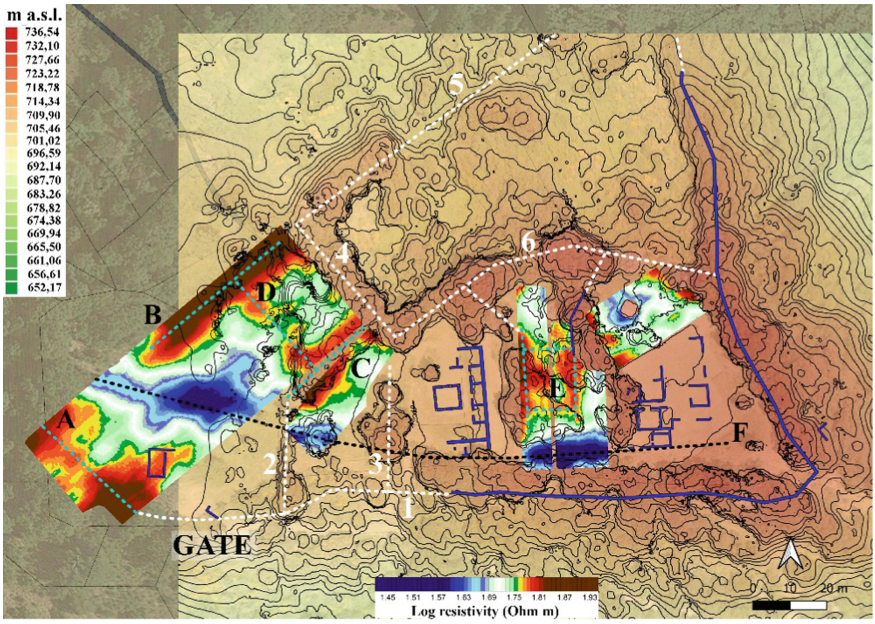


Fig. 7. Interpretation of results considering the DEM-derived contour model, the resistivity map relative to the frequency of 15 kHz and previously recorded remains.

The analysis of resistivity data provides more evidence for these spatial projections. High resistivity readings indicate a symmetrical and structured distribution with respect to the known architectural remnants and the probable terrace walls. Such regularity is frequently associated with compacted anthropogenic layers seen in highly conductive sedimentary matrices or stone-built buildings. The suggested city wall restoration (section 1) at point A (Fig. 7) is represented by a distinct resistive band. Its geometric continuity strongly suggests that this anomaly may be connected to the underground extension of the defensive construction.

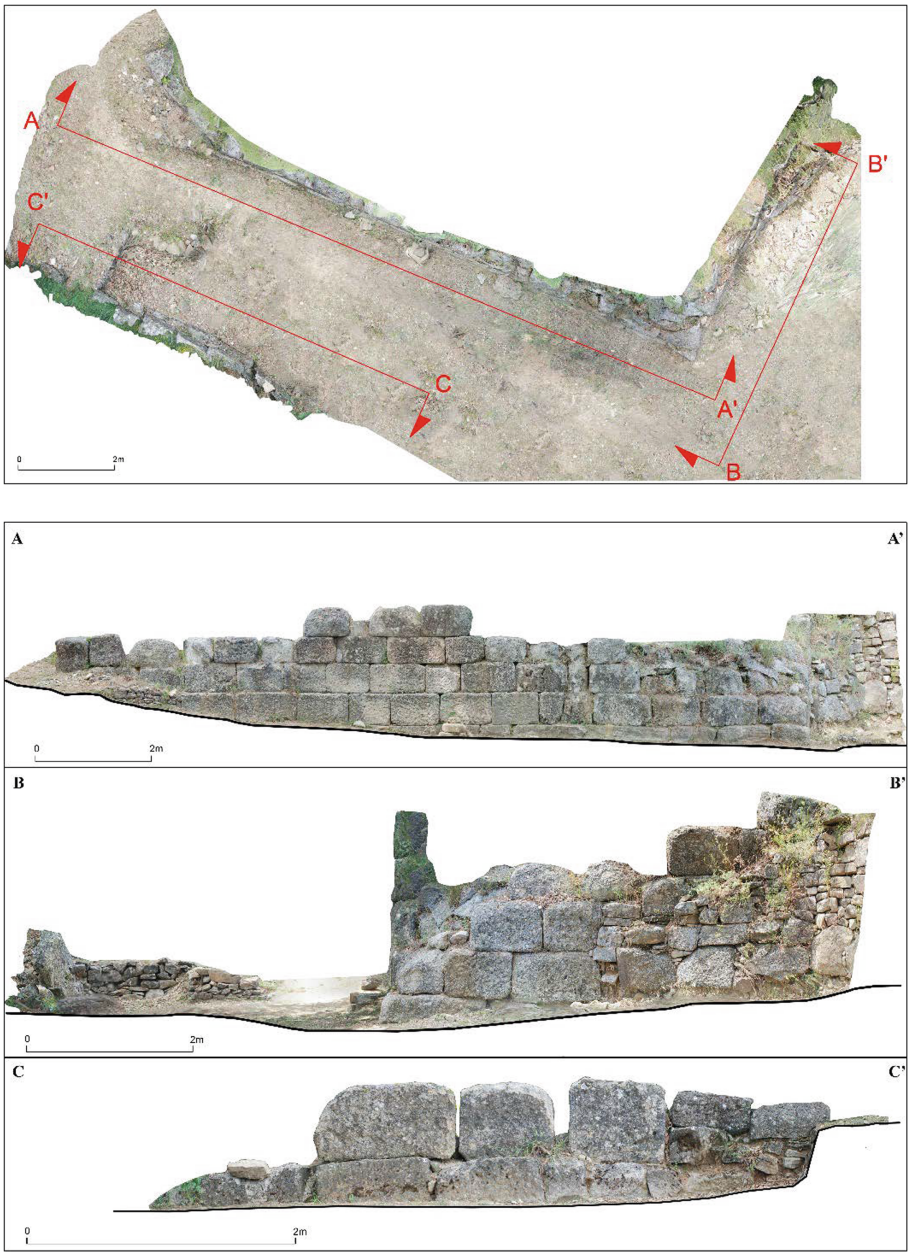


Fig. 8. Orthophoto of the gate (top) with indication of the underlaying architectural elevations A-A′, B-B′ and C-C′ (bottom).

Anomalies B, C, and D designate another possibly significant zone. When combined with the suggested terracing indicated by number 4, these resistive characteristics produce a rectangular area whose purpose and chronology are still unknown. The geometric regularity of this layout may indicate the presence of a building complex, an enclosed functional space or a courtyard. Remarkably, resistive segments B and C show structural integration with the broader architectural concept by converging in the direction of terrace sections 5 and 6.

A possible quadrangular underlying structure is suggested by the spatial distribution of resistivity peaks in the center of the excavation region (Anomaly E). This peculiarity appears to be partially connected to a wall that has already been found during previous archaeological digs along its northern side. This feature could represent an extra room or an extension within the settlement's constructed fabric.

Conversely, areas with low resistivity values, shown by a black dotted line and the symbol F, correspond to an underground pipeline that has previously been found. The lower resistivity in this area is consistent with increased moisture retention and conductive backfill materials, which are frequently associated with such infrastructural components. Evidence from an eastern archaeological test trench and the inferred route from the geophysical survey match, demonstrating the interpretive reliability of the geophysical dataset.

Overall, the combination of high-resolution aerial photogrammetry and non-invasive geophysical prospection has enabled a comprehensive spatial reconstruction of the Castello site. In addition to synthesizing current knowledge, the resulting multi-layered territorial information system generates new hypotheses about the terracing systems, the settlement structure and the defensive architecture. This methodological synergy provides a strong scientific basis for future excavation plans, conservation initiatives and targeted archaeological studies.

4 Conclusions

An improvement in understanding of the Castello archaeological site has been achieved by the systematic integration of non-invasive techniques, such as geophysical prospections and both aerial and terrestrial photogrammetry. This has helped to partially overcome the interpretive limitations inherent in traditional excavation-based investigations alone especially in situations where above-ground structures are discontinuously preserved and there is dense vegetation cover. Anomalies associated with the existence of buried masonry structures, potential manmade terracing systems and other architectural features not yet excavated were supposed through geophysical investigations. One of the main unanswered questions arising from earlier research was addressed by the spatial distribution of resistivity anomalies, which offered credible evidence regarding the continuation of the defensive circuit in the northern and southwestern sectors. Additionally, the identification of regular planimetric layouts within the defended region indicates the presence of additional ordered areas, the functional interpretation of which will necessitate focused stratigraphic verification.

Simultaneously, high-resolution georeferenced three-dimensional models and a Digital Elevation Model (DEM) capable of precisely depicting site morphology and altimetric relationships between archaeological structures and the surrounding landscape

were produced by photogrammetry, which was implemented through both remotely piloted aircraft systems (RPAS) and ground-based image acquisition. More accurate topographic contextualization of archaeological material was obtained by the integration of these outputs into a Geographical Information System, which allowed for extensive data management.

From a historical-archaeological standpoint, the findings help to clarify the site's urban layout by emphasizing a settlement planning approach that is in line with the territorial control and the defensive roles typical of fortified centers in the Lucanian-Calabrian region in the third century B.C. Furthermore, the integrated research indicates that the settlement's internal organization was more intricate than previously recorded, maybe divided into regions with distinct functions.

The insights generated by this study offer a strong scientific foundation for organizing upcoming research projects, directing excavation efforts toward regions with greater archaeological value while reducing the damaging effects of field operations. Furthermore, the created three-dimensional and cartographic documentation serves as a strategic tool for public distribution, site management and heritage protection, supporting scientific communication and conservation planning.

In summary, the Castello site research demonstrates how geomatic and geophysical technologies can not only support but also significantly improve the interpretative and operational strategies of modern archaeological research, confirming the efficacy of an integrated and interdisciplinary approach in the study of ancient fortified landscapes.

Acknowledgments. This study was carried out in accordance with the agreement between the Municipality of Cersosimo and the Department of Humanities, Social Sciences, and Education at the University of Molise for "Scientific and operational collaboration for the implementation of integrated non-invasive diagnostic investigations in the Castello Archaeological Park, aimed at guiding and supporting strategic decisions for the restoration, improved use and enhancement of the monument".

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

Declaration The authors used *QuillBot Paraphraser* [https://quillbot.com/] to improve English language during the preparation of this work. After using this tool, the authors reviewed and edited the content as needed and takes full responsibility for the content of the publication.

References

1. Bianco, S., Tarlano, F.: Valsinni Monte Coppolo. Le Fortificazioni Italiche Del IV Sec. a.C. In: Calalogo della mostra. Soprintendenza per i Beni Archeologici della Basilicata, Valsinni, Italia (2011)
2. Lacava, M.: Cersosimo. Degli avanzi di un antichissimo abitato in contrada Castello presso Cersosimo nel circondario di Lagonegro in Basilicata. NSA. **1889**, 88–90 (1889)
3. Lacava, M.: Topografia e storia di Metaponto, pp. 346–350. Morano, Napoli (1891)
4. De La Genière, J.: Aspetti e problemi dell'archeologia del mondo indigeno. In: ACT XI, pp. 258–259 (1971)
5. De La Genière, J.: Siti: Carancino – Crotone. Bibliografia topografica della colonizzazione greca in Italia e nelle Isole Tirreniche, pp. 248–250 (1987)

6. Osanna, M.: Paesaggi agrari e organizzazione del territorio in Lucania tra IV e III sec. a.C. Boll. Archeol. Online. **A/A1/4**, 17–31 (2010)
7. Piroddi, L. et al.: Imaging cultural heritage at different scales: part II, the Meso-scale (sites). Remote Sens. **17**(4), 598 (2025)
8. Cozzolino, M., Di Giovanni, E., Mauriello, P., Piro, S., Zamuner, D.: Geophysical Methods for Cultural Heritage Management Springer Geophysics Series. Springer, Cham, Switzerland (2018)
9. Oswin, J.: A Field Guide to Geophysics in Archaeology. Springer Science & Business Media, Berlin, Germany (2009)
10. Witten, A.: Handbook of Geophysics and Archaeology. Routledge, London, UK (2017)
11. Schmidt, A.: Geophysical Data in Archaeology: a Guide to Good Practice. Oxbow Books, Oxford, UK (2013)
12. Gaffney, V. et al.: Durrington walls and the Stonehenge hidden landscape project 2010–2016. Archaeol. Prospect. **25**, 255–269 (2018)
13. Urban, T.M. et al.: Magnetic detection of archaeological hearths in Alaska: a tool for investigating the full span of human presence at the gateway to North America. Quat. Sci. Rev. **211**, 73–92 (2019)
14. Fitton, T., Contreras, D.A., Gidna, A.O., Mabulla, A.Z., Prendergast, M.E., Grillo, K.M.: Detecting and mapping the 'ephemeral': Magnetometric survey of a pastoral neolithic settlement at Luxmanda Tanzania. Antiquity. **96**, 298–318 (2022)
15. Conyers, L.B., Ernenwein, E.G., Grealy, M., Lowe, K.M.: Electromagnetic conductivity mapping for site prediction in Meandering River floodplains. Archaeol. Prospect. **15**, 81–91 (2008)
16. De Smedt, P. et al.: Exploring the potential of multi-receiver EMI survey for geoarchaeological prospection: a 90 ha dataset. Geoderma. **199**, 30–36 (2013)
17. Cozzolino, M., Mauriello, P., Patella, D.: The extended data-adaptive probability-based electrical resistivity tomography inversion method (E-PERTI) for the characterization of the buried ditch of the ancient Egnazia Puglia, Italy. Appl. Sci. **12**, 2690 (2022)
18. Papadopoulos, N.G., Yi, M.J., Kim, J.H., Tsourlos, P., Tsokas, G.N.: Geophysical investigation of tumuli by means of surface 3D electrical resistivity tomography. J. Appl. Geophys. **70**, 192–205 (2010)
19. Al-Saadi, O.S., Schmidt, V., Becken, M., Fritsch, T.: Very-high-resolution electrical resistivity imaging of buried foundations of a Roman villa near Nonnweiler Germany. Archaeol. Prospect. **25**, 209–218 (2018)
20. Cozzolino, M., Di Giovanni, E., Gentile, V., Mauriello, P., Pizzano, N.: Ground-penetrating radar survey for the study of the Church of Saint Cosma in Helerito (Tagliacozzo, L'Aquila, Italy). Geosciences. **10**, 244 (2020)
21. Booth, A.D., Szpakowska, K., Pischikova, E., Griffin, K.: Structure of an ancient Egyptian tomb inferred from ground-penetrating radar imaging of deflected overburden horizons. Archaeol. Prospect. **22**, 33–44 (2015)
22. Urban, T.M., Murray, C.A., Vella, C., Lahikainen, A.: Ground-penetrating radar survey on the island of Pantelleria (Italy) reveals an ancient architectural complex with likely Punic and Roman components. J. Appl. Geophys. **123**, 164–169 (2015)
23. Schmidt, A., Linford, P., Lindford, N., Gaffney, C., David, A. (eds.): EAC Guidelines for the Use of Geophysics in Archaeology. Archaeolingua Press, Budapest (2015)
24. Reynolds, J.M.: An Introduction to Applied and Environmental Geophysics. John Wiley & Sons, Chichester (2011)
25. Profiler EMP-400: GSSI Geophysical Survey Systems, (2026). www.geophysical.com. Last Accessed 30 Jan 2026
26. Dipanda, A., Woo, S.: Towards a real-time 3D shape reconstruction using a structured light system. Pattern Recogn. **38**(10), 1632–1650 (2005)

27. Prakoonwit, S., Benjamin, R.: 3D surface point and wireframe reconstruction from multiview photographic images. Image Vis. Comput. **25**(9), 1509–1518 (2007)
28. Park, J.S.: Interactive 3D reconstruction from multiple images: a primitive-based approach. Pattern Recogn. Lett. **26**(16), 2558–2257 (2005)
29. Angelini, A., Cozzolino, M., Gabrielli, R., Gentile, V., Mauriello, P.: Three-dimensional Modeling and non-invasive diagnosis of a huge and complex heritage building: the patriarchal basilica of Santa Maria Assunta in Aquileia (Udine, Italy). Remote Sens. **15**, 2386 (2023)
30. Colomina, I., Molina, P.: Unmanned aerial systems for photogrammetry and remote sensing: a review. ISPRS J. Photogramm. Remote Sens. **92**, 79–97 (2014)
31. Murtiyoso, A., Grussenmeyer, P.: Documentation of heritage buildings using close-range uav images: dense matching issues, comparison and case studies. Photogramm. Rec. **32**, 206–229 (2017)
32. Marín-Buzón, C., Pérez-Romero, A., López-Castro, J.L., Ben Jerbania, I., Manzano-Agugliaro, F.: Photogrammetry as a new scientific tool in archaeology: worldwide research trends. Sustainability. **13**, 5319 (2021)
33. Verhoeven, G.: Taking computer vision aloft—archaeological three-dimensional reconstructions from aerial photographs with photoscan. Archaeol. Prospect. **18**, 67–73 (2011)
34. Hwang, J.T., Ting-Chen, C.: 3D building reconstruction by Multiview images and the integrated application with augmented reality the international archives of photogrammetry. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. **41**, 1235 (2016)
35. Kholil, M., Ismanto, I., Fu'Ad, M.N.: 3D reconstruction using structure from motion (SFM) algorithm and multi view stereo (MVS) based on computer vision. In: IOP Conference Series: Materials Science and Engineering, vol. 1073, p. 012066. IOP Publishing (2021)
36. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. Int. J. Comput. Vis. **47**(1–3), 7–42 (2002)
37. Okutomi, M., Kanade, T.: A multiple-baseline stereo. IEEE Trans. Pattern Anal. Mach. Intell. **15**(4), 353–363 (1993)
38. Agisoft Metashape User Manual: Standard Edition, Version 2.0, (2026). https://www.agisoft.com/pdf/metashape_2_0_en.pdf. Last Accessed 9 Mar 2026

Assessment of Low-Cost GNSS Receiver Performance in Urban Environments

Giuseppina Vacca(✉), Enrica Vecchi, and Antonio Pala

Department of Civil, Environmental Engineering and Architecture, University of Cagliari, Cagliari 09123, Italy
{giuseppina.vacca,vaccag}@unica.it

Abstract. Global Navigation Satellite System (GNSS) enables positioning accuracies ranging from meter-level to sub-centimeter, traditionally achieved using geodetic-grade receivers. However, the high cost of such instruments still limits their accessibility in many practical applications. In this context, the recent development of multi-frequency, multi-constellation low-cost GNSS receivers has opened new perspectives for cost-effective high-precision positioning, although their performance in challenging environments remains to be fully assessed. This study evaluates the performance of a multi-frequency low-cost (LC) GNSS receiver (u-blox ZED-F9P with ANN-MB-00 patch antenna) in comparison with a geodetic-grade device (GD), Trimble R12. Experiments were conducted in two representative urban environments: a semi-open area (ST1) and an obstructed area characterized by buildings and vegetation (ST2). Both static relative positioning and Network Real-Time Kinematic (NRTK) techniques were investigated, using the geodetic receiver as a reference. Signal quality was assessed in terms of carrier-to-noise ratio (C/N_0) and multipath (MP) errors. In static positioning, millimeter-level agreement with the geodetic solution is achieved in horizontal components under favorable conditions, while errors increase to the centimeter level in obstructed environments. The NRTK results confirm this behavior, with sub-centimeter horizontal dispersion in semi-open conditions and a degradation up to the centimeter level in the obstructed environment. The vertical component consistently exhibits the largest errors, with systematic biases and higher variability, confirming its greater sensitivity to environmental effects. Overall, the results highlight the potential of multi-frequency LC GNSS receivers for high-precision applications under favorable conditions, while emphasizing their sensitivity to multipath and signal obstruction in urban environments.

Keywords: GNSS · low-cost receivers · u-blox · NRTK · urban surveying

1 Introduction

Global Navigation Satellite System (GNSS) technology has been widely used for many years across a broad range of applications, from navigation to high-precision static positioning, with achievable accuracy ranging from a few meters down to the sub-centimeter levels [1].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 86–103, 2027.
https://doi.org/10.1007/978-3-032-30539-8_6

In particular, the concept of high-precision GNSS positioning, for both post-processing and real-time applications, has traditionally been associated with geodetic-grade receivers, which are characterized by high performance, robustness, and well-established reliability. However, despite their widespread availability on the market, the cost of such equipment remains high, limiting its accessibility for many users, especially in application fields where cost-effectiveness is a primary concern.

In this context, low-cost receivers have been progressively developed, initially based on single-frequency modules [2]. As expected, the lack of dual-frequency observations at that stage constrained their performance, both in terms of accuracy and operational capabilities [3].

In parallel, recent years have undergone the rapid expansion and full operational capability of multiple GNSS constellations, resulting in a substantial increase in the number of available satellites. At present, more than 110 satellites are operational across four different constellations, 99 of which are in the Medium Earth Orbit (MEO): GPS (29), GLONASS (24), Galileo (22), and BeiDou (24). Consequently, multi-constellation positioning based on a large number of simultaneously visible satellites has rapidly become a common scenario [4].

This technological evolution has significantly driven the development of GNSS instruments, particularly in the low-cost market. Indeed, modern low-cost devices now support multi-frequency, multi-constellation observations and are often equipped with Real-Time Kinematic (RTK) and Network RTK (NRTK) capabilities. These features have made low-cost instruments suitable for precise positioning in a wide range of geodetic and surveying applications [1]. As a result, it is now feasible to assemble low-cost GNSS systems based on OEM (Original Equipment Manufacturer) modules, offering a cost-effective alternative to commercial geodetic receivers currently available on the market. The price difference is notable, with low-cost receivers typically costing a few hundred euros, compared to over 10 k–15 k euros for geodetic devices.

Another key distinction between low-cost GNSS receivers and geodetic-grade instruments lies in the way they are delivered to users. Low-cost receivers are typically provided as modular components that require ad hoc assembly and integration to form a fully operational setup, while geodetic-grade receivers are generally supplied as complete, ready-to-use solutions.

Several studies have investigated the design and performance of multi-frequency low-cost GNSS (MF-LC GNSS) receivers. Commercial devices such as u-blox, Emlid, Sky-Traq, Swift, and Piksi Multi have demonstrated the capability of achieving centimeter-level accuracy, particularly in open and rural environments [3, 5, 6]. In such environments, relative positioning solutions can achieve optimal accuracies, reaching up to 1 mm in the horizontal components and 3 mm in height [7, 8]. However, the scientific literature addressing GNSS performance in urban or densely vegetated environments remains relatively limited [9].

The present study aims to evaluate the performance of a multi-frequency (MF) low-cost (LC) GNSS receiver under urban and obstructed conditions, with particular emphasis on its comparison with a geodetic-grade device (GD). Specifically, the u-blox ZED-F9P MF-LC receiver, equipped with the u-blox ANN-MB-00 antenna, was adopted as the

low-cost solution, while a Trimble R12 receiver was used as a reference to assess positioning accuracy and overall performance. The low-cost receiver costs approximately 400 euros, while the geodetic-grade receiver exceeds 15,000 euros.

The remainder of the paper is organized as follows: Sect. 2 presents the instrumentation, study area, and experimental setup; Sect. 3 reports the results; and Sect. 4 provides a discussion of the results and the main conclusions.

2 Materials and Methods

2.1 GNSS Devices

Low-Cost Receiver

In this study, a low-cost GNSS receiver, the ZED-F9P-02B (hereafter referred to as "LC"), was employed: This device integrates multi-band Real-Time Kinematic (RTK) technology, enabling centimeter-level positioning accuracy under favorable conditions. The very compact size of the module (17.0 × 22.0 × 2.4 mm) makes it suitable for several applications. The ZED-F9P-02B supports multi-constellation and multi-frequency signal tracking, including GPS (L1 C/A, L2C), GLONASS (L1OF, L2OF), Galileo (E1, E5b), and BeiDou (B1I, B2I). This multi-band capability improves robustness against multipath and enhances ambiguity resolution performance compared to single-frequency receivers. Detailed specifications are available at [10].

The LC receiver was combined with a compact and versatile multi-band patch antenna, the ANN-MB-00-00, also produced by u-blox, as shown in Fig. 1.

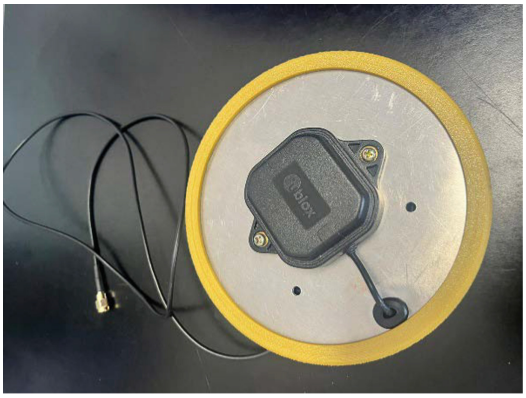


Fig. 1. u-blox ANN-MB-00-00 multi-band antenna mounted on a 3D-printed ground plane.

To ensure optimal signal reception and proper functionality, the antenna was mounted on a custom 3D-printed ground plane, providing a stable reference surface and minimizing potential interference from surrounding components (Fig. 1). This antenna provides stable performance across multi-frequency and is well suited for integration with the ZED-F9P-02B module. Additional specifications are available at [11].

Geodetic Receiver

For validation purposes, a geodetic-grade GNSS receiver (Trimble R12), hereafter referred to as "GD", was used to define the ground truth. The Trimble R12 is a high-performance multi-frequency, multi-constellation receiver equipped with 672 channels, capable of tracking GPS, GLONASS, Galileo, BeiDou, and QZSS signals. It is designed for professional surveying applications and provides robust RTK performances, and enhanced resistance to challenging environments such as multipath and signal obstruction [12].

2.2 Study Area

The study area is located within the Faculty of Engineering and Architecture of the University of Cagliari (Sardinia, Italy). Two test environments representative of typical urban surveying conditions were selected (Fig. 2), reflecting contexts in which completely open-sky scenarios are rarely encountered. These environments are hereafter referred to as "ST1" and "ST2".

- ST1: semi-open environment, with limited obstructions from surrounding buildings or vegetation and relatively good sky visibility. The survey point in ST1 is located on a flat paved courtyard, surrounded by Faculty buildings (approximately 5–10 m in height) and sparse tall trees. Both types of obstacles are sufficiently distant from the observation point to preserve adequate sky visibility.
- ST2: obstructed area, affected by the combined presence of buildings and dense tall trees. The survey point in ST2 is located along a flat passageway crossing a tree-lined courtyard. In this case, the primary source of obstruction is represented by the surrounding vegetation, while Faculty buildings (with heights comparable to those in ST1) are also present at greater distances from the observation point.

At each site, two nearby survey points (denoted as ST' and ST") were established. These points were positioned as close as possible to each other while still allowing sufficient space for tripod installation (approximately 70 cm separation), as shown in Fig. 3. This configuration ensures comparable satellite geometry and visibility conditions at both points, thus minimizing differences related to instruments placement.

2.3 Quality Control of GNSS Observations

Low-cost GNSS receivers generally exhibit lower performance compared to geodetic-grade receivers, primarily due to differences in hardware design [3, 13, 14]. To assess GNSS observation quality, the carrier-to-noise ratio (C/N_0) and multipath error (MP) were considered as the primary metrics. Low-cost receivers typically exhibit lower C/N_0 values due to reduced signal tracking performance [5, 14]. This effect can be assessed by analyzing corresponding observations collected using both LC and GD receivers under semi-open sky (ST1) and obstructed (ST2) conditions. For geodetic receivers, C/N_0 values generally range between 35–45 dB-Hz, decreasing below 30 dB-Hz at low elevation angles [15].



Fig. 2. Map of the survey points, showing ST1 and ST2, and the UCAG permanent GNSS station located on the rooftop of the Faculty of Engineering building at the University of Cagliari. Grid coordinates are aligned to WGS84 system (EPSG:4326), basemap is derived from Google Satellite imagery.

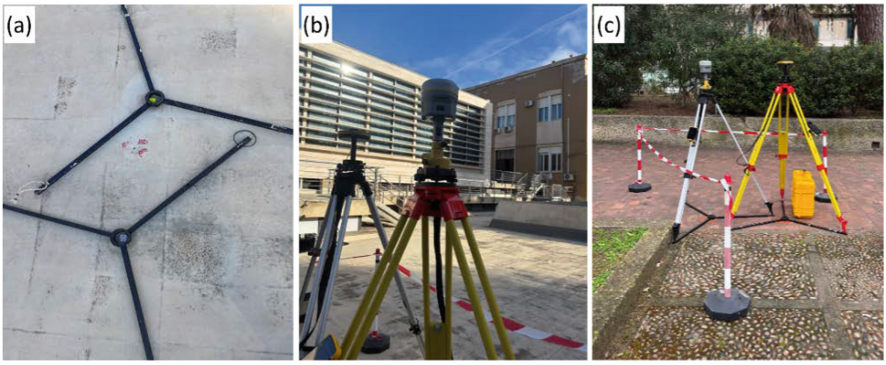


Fig. 3. Configuration of the two nearby survey points (ST' and ST"), illustrated for the ST1 environment (a), and setup of the two GNSS receivers operating at ST1 (b) and ST2 (c).

Multipath errors, meaning the effects occurring when the GNSS signal received by the antenna results from the superposition of direct and reflected (or refracted) components, are strongly environment-dependent due to the presence of surrounding objects (buildings, terrain, and vegetation) [16]. Moreover, since multipath implies that both direct and reflected signals enter the receiver through the antenna, mitigation strategies often focus on antenna design, which represent one of the main parameters affecting its final effect [17, 18]. For instance, specialized antennas such as right-handed circularly polarized or Choke-Ring designs are commonly employed to mitigate multipath effects [19]. Low-cost antennas, by contrast, are generally more susceptible due to their reduced

ability to discriminate against reflected signals, also because they are commonly provided without dedicated design support. The issue is particularly pronounced in urban and vegetated environments, where dense reflective surfaces and foliage increase signal scattering, ultimately leading to degradation in positioning accuracy [20].

The chosen configuration of the survey points (see Sect. 2.2), with LC and GD receivers installed approximately 70 cm apart, allowed the evaluation of these factors under identical observation conditions.

2.4 GNSS Positioning and Data Processing

Static Sessions

The positioning accuracy of the low-cost (LC) receiver was assessed by comparison with the geodetic-grade (GD) receiver operating at the same survey point. Thus, for each environment (semi-open sky and obstructed area), two independent datasets were collected, with the GD providing the ground truth. Measurement sessions were conducted one day apart to ensure similar environmental and satellite conditions. This experimental design was necessary because each survey point had to be independently validated using the GD receiver.

The acquisition scheme is summarized as follows:

Semi-open environment - ST1

- Session 1A
  - Point ST1' → GD
  - Point ST1" → LC
- Session 1B (instrument swap, one day apart):
  - Point ST1' → LC
  - Point ST1" → GD

Obstructed area environment - ST2

- Session 2A:
  - Point ST2' → GD
  - Point ST2" → LC
- Session 2B (instrument swap, one day apart):
  - Point ST2' → LC
  - Point ST2" → GD

A common observation interval shared by all sessions across both environments (ST1 and ST2) was selected, resulting in an overlapping duration of approximately 4 h. Although the first two sessions were originally longer, they were truncated to ensure

Table 1. GNSS acquisition sessions at ST1 and ST2, showing session IDs, and overlapping observation time intervals (UTC), with start and stop times of each acquisition.

Environment	Session	Observation time	
		Start (UTC)	Stop (UTC)
Semi-open ST1	1A – day 1	08:29:46	15:10:36
	1B – day 2	08:31:22	14:44:53
Obstructed ST2	2A – day 1	08:35:40	12:35:51
	2B – day 2	08:22:12	12:24:16

consistency and comparability across all datasets under similar observational conditions (Table 1).

Positioning was performed using baselines computed with respect to the UCAG (UCAG00ITA) permanent GNSS station, located on the rooftop of the Engineering building of the University of Cagliari. The UCAG station, operated by the Italian Space Agency (ASI) [21], has been included in the European Permanent Network (EPN) since August 2015 and is classified as Class A station. As such, it provides a stable and reliable reference for high-precision GNSS positioning and baseline analysis.

Due to the short distance between the survey points and the reference station (<1 km), all baselines (ST1'- UCAG, ST1"- UCAG, ST2'- UCAG, ST2"- UCAG) can be classified as short baselines.

Data processing was performed using the open-source software RTKLIB-EX (version 2.5.0) [22], with parameters summarized in Table 2. Coordinates (E, N, h) were computed in the official national reference system ETRS89/ETRF2000, and positioning errors were assessed by comparing the LC-derived coordinates with the GD reference solution.

Table 2. Main parameters adopted for static GNSS processing using RTKLIB.

Parameter	Value
Signal frequencies	Dual frequency
GNSS Constellations	GPS, GLONASS, Galileo, BeiDou, SBAS
Elevation mask	15 deg
Sampling interval	15 s
Tropospheric model	Saastamoinen
Ionospheric model	Broadcast
Ephemerides	Broadcast

NRTK Sessions

Additional tests were conducted using the LC receiver in NRTK (Network Real-time

Kinematic) mode. Corrections were provided by the Sardinian permanent GNSS network SARNET [23], ensuring alignment with the ETRF2000-UTM32N coordinate system (EPSG: 6707).

In particular, the performance of the LC receiver in NRTK mode was evaluated through repeated short-duration sessions designed to simulate real operational conditions. Measurements were carried out on two consecutive days, at one point for each environment, following the same acquisition strategy to ensure consistency and comparability between the two scenarios.

The receiver remained continuously active during the observation period, while individual points were acquired at regular 15-min intervals, in order to assess temporal variability under changing satellite geometry conditions. The same time window adopted for the static sessions was used for consistency (approximately between 09:00 and 13:00 UTC), with 5 s observations, resulting in 17 independent measurements for each environment. The antenna was mounted on a 2 m survey pole, properly centered over the point and stabilized using a bipod (Fig. 4), ensuring adequate centering and leveling stability throughout the entire observation period.



Fig. 4. GNSS antenna setup with bipod support during the NRTK measurements (ST1).

3 Results

3.1 Carrier-to-Noise Ratio (C/N_0) Analysis

This section presents the analysis of the carrier-to-noise ratio (C/N_0) as a primary indicator of signal quality, comparing the low-cost (LC) and geodetic-grade (GD) receivers under the two selected different environmental conditions. The comparison is performed

between the two frequency bands (L1 and L2), considering all tracked constellations together. In the L1 band, signals such as L1 C/A, L1OF, E1, and B1I are included, while in the L2 band signals such as L2C, L2OF, E5b, and B2I are considered.

For clarity, results are presented for only one representative point per session, reflecting the corresponding environmental conditions (semi-open and obstructed). Figures 5 and 6 depict the Carrier-to-noise ratio (C/N_0) as a function of satellite elevation angle for the LC and GD receivers, respectively, based on RTKLIB processing, while Table 3 summarizes the main results.

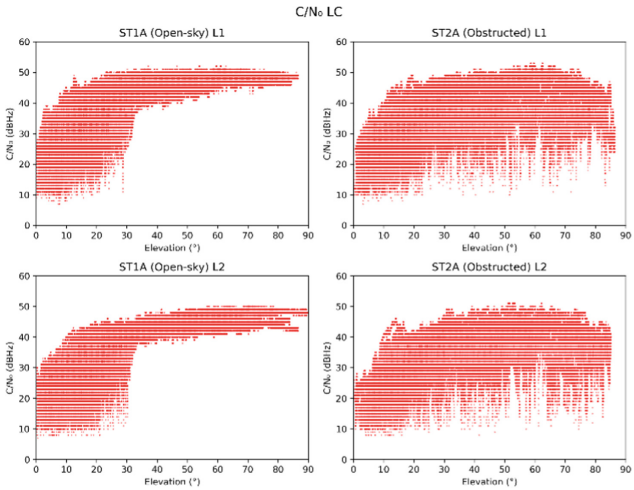


Fig. 5. C/N_0 for LC receiver as a function of satellite elevation angle.

Observing the C/N_0 graphs in Figs. 5 and 6 as a function of satellite elevation angle for LC and GD receivers, respectively, and considering the statistical summary in Table 3, the following trends can be identified:

- **LC vs GD:** the geodetic receiver consistently exhibits higher SNR values compared with the low-cost receiver, indicating superior signal tracking performance and measurement stability. The LC receiver shows a higher variability (higher Std value), resulting in more sparse distribution.
- **L1 vs L2:** for both receivers, L1 signals show slightly higher values compared to L2, confirming the generally better signal quality and robustness of the L1 frequency, however, this small variation is not clearly noticeable in the distributions and becomes evident primarily from the statistical results.
- **ST1 vs ST2:** a clear degradation in C/N_0 is observed in the obstructed environment (ST2), with a more pronounced reduction for the low-cost receiver, highlighting its higher sensitivity to signal attenuation. In particular, the mean value decreases by approximately 7.9% for the geodetic receiver and 8.8% for the low-cost receiver from ST1 to ST2.

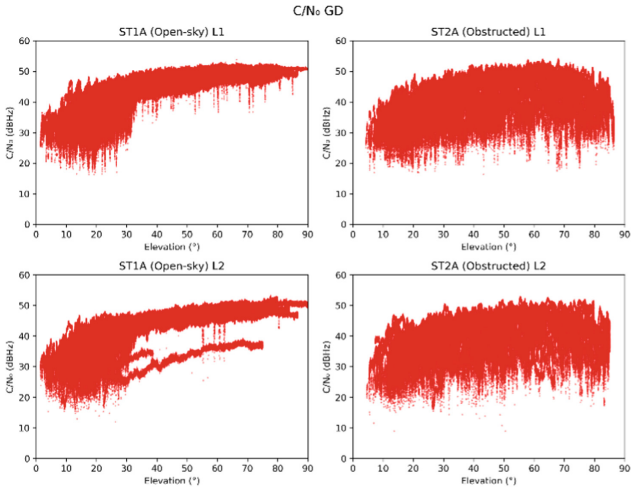


Fig. 6. C/N_0 for GD receiver as a function of satellite elevation angle.

Table 3. Statistical summary of C/N_0 values for the LC and GD receivers, reported for one session in ST1 (S1A) and one in ST2 (S2A), for both frequencies. Mean, median, and standard deviation (Std) are provided; all values are expressed in dB-Hz.

Receiver	Session	Frequency	Mean	Median	Std
GD	S1A	L1	44.5	46.4	6.1
	S1A	L2	42.7	44.7	6.1
	S2A	L1	41.0	41.7	6.4
	S2A	L2	39.9	40.6	6.0
LC	S1A	L1	41.9	45.0	7.6
	S1A	L2	38.7	42.0	8.0
	S2A	L1	38.2	39.0	8.1
	S2A	L2	36.0	38.0	7.8

3.2 Multipath Effects

The multipath error analysis is reported in this subsection, with the aim of evaluating the impact of antenna quality and surrounding environment on signal reflections and measurement distortions. As in the previous case, only one representative point from each session is presented. Figures 7 and 8 show the multipath errors for the LC and GD receivers, respectively, plotted against the satellite elevation angle, and Table 4 summarizes the main results.

Based on the multipath distributions shown in Figs. 7 and 8 for the LC and GD receivers, together with the statistical summary reported in Table 4, the main observed trends are summarized as follows:

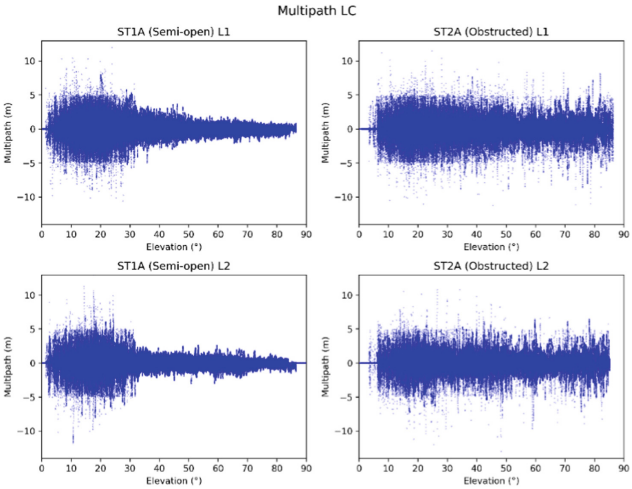


Fig. 7. Multipath LC.

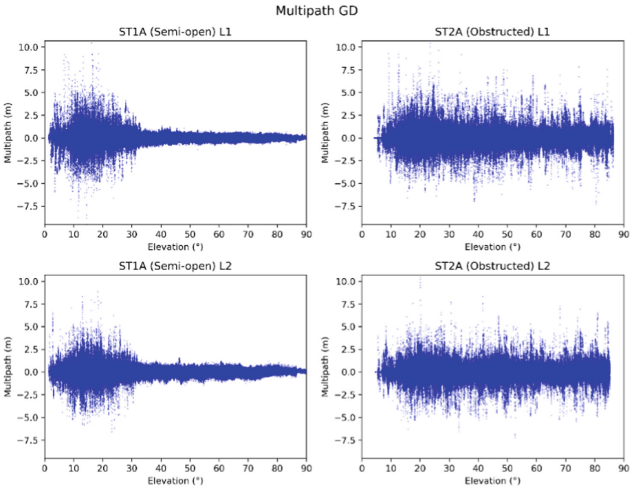


Fig. 8. Multipath GD.

- **LC vs GD:** the LC receiver exhibits significantly higher (about two times) multipath levels compared with the geodetic receiver, primarily due to the lower quality of the antenna system which is known to be the major factor affecting this parameter [18].
- **L1 vs L2:** minor differences are observed between L1 and L2 under the same environmental conditions.
- **ST1 vs ST2:** multipath effects increase substantially in the obstructed environment, confirming the strong influence of surrounding obstacles such as vegetation and buildings. The trend is clearly visible both in the graphs and in the statistical results. In

Table 4. Statistical summary of MP values for the LC and GD receivers, reported for one session in ST1 (S1A) and one in ST2 (S2A), for both L1 and L2 frequencies. Standard deviations (Std) are provided, with values expressed in meters.

Receiver	Session	L1 Std	L2 Std
GD	S1A	0.46 m	0.47 m
GD	S2A	0.82 m	0.80 m
LC	S1A	0.88 m	0.87 m
LC	S2A	1.21 m	1.14 m

particular, the standard deviation increases by approximately 70–78% for the geodetic receiver and 31–38% for the low-cost receiver from ST1 to ST2.

It is worth noting that multipath levels for a GD receiver are typically lower in fully open-sky conditions than those observed at ST1. Indeed, ST1 is a semi-open environment, where nearby buildings and vegetation introduce residual reflections (Figs. 2 and 3).

3.3 Positioning Accuracy

Static Positioning

Static positioning performance is assessed by comparing the coordinates estimated by the LC receiver with those obtained from the GD, which is used as reference. The influence of environmental conditions is also evaluated by comparing results in ST1 and ST2.

For each dataset, the static solution is computed as the mean of all available epochs over the entire observation period (approximately 4 h), using RTKLIB software with the parameters listed in Table 2. Based on these positions, coordinate differences between LC and GD solutions are calculated for the Easting (E), Northing (N), and ellipsoidal height (h) components. These differences represent the positioning bias and are used as indicators of accuracy with respect to the reference solution. In addition, the solution quality (Q) and the number of tracked satellites are examined to further assess the reliability of the positioning results under the two environmental conditions. A three-dimensional error metric is computed as the Euclidean distance between the LC and GD coordinates, representing the overall positioning error derived from the averaged solutions.

Table 5 summarizes the comparison between the GD and LC solutions for the two environments. In ST1, the LC-GD differences are at the millimeter level in the horizontal components, while the vertical component shows a difference of about 1.3 cm. The resulting horizontal error is approximately 3 mm, while the 3D error remains at the centimeter level, confirming a high level of agreement between the two receivers under favorable conditions.

In both environments, fixed solutions ($Q = 1$) are achieved by both receivers, despite differences in the number of tracked satellites. In ST2, the discrepancies increase, particularly in the Northing component (1.6 cm), while the Easting difference remains limited

Table 5. Comparison of LC and GD static solutions in ST1 and ST2. The quality indicator (Q), number of satellites (NS), coordinate differences (LC - GD, in meters), and horizontal and 3D errors are reported.

Env	Q_GD	NS_GD	Q_LC	NS_LC	dE (m)	dN (m)	dh (m)	error hor	error 3D
ST1	1	17	1	17	−0.002	0.002	−0.013	0.003	0.013
ST2	1	13	1	18	0.002	0.016	−0.031	0.016	0.035

(2 mm) and comparable to the values observed in ST1. The vertical component shows a larger bias of about 3.1 cm. Consequently, the horizontal error rises to approximately 1.6 cm, and the 3D error reaches about 3.5 cm, indicating a degradation in positioning performance. These results suggest that, although fixed ambiguity solutions are maintained in both environments, positioning accuracy decreases under more challenging conditions, likely due to increased multipath effects and signal obstructions, with a more pronounced impact on the vertical component.

NRTK Performance

The positioning performance in NRTK mode was evaluated in terms of temporal stability, defined as the repeatability of the estimated coordinates over time and their sensitivity to variations in satellite geometry. The assessment was conducted using repeated short-duration observations acquired at regular time intervals, and the results were then compared with those obtained under static conditions. Furthermore, measurements collected in the two environments (ST1, ST2) were analyzed to investigate the influence of surrounding conditions on positioning performance.

Figure 9 illustrates the time series of the three position components during the NRTK session in ST1, together with their corresponding mean values and standard deviations. The horizontal components exhibit a good stability throughout the observation period, with standard deviations of 3.5 mm and 5.8 mm, for the Easting and Northing components, respectively. This results in a combined horizontal dispersion at the sub-centimeter level (less than 7 mm), with a slightly larger contribution from the Northing component.

The residuals of each epoch with respect to the corresponding mean value are generally contained within approximately ±0.01 m, indicating a stable and well-converged solution over time. The vertical component shows higher variability, with a standard deviation of approximately 1.4 cm over the 17 solutions considered. Consistently, the height residuals exhibit a wider dispersion compared to the horizontal components.

A notable outlier is observed in the vertical component, reaching approximately +0.042 m relative to the mean position, around 12:00 UTC. This deviation may be attributed to a temporary degradation in satellite geometry or multipath effects. Overall, this behavior is expected, as the vertical component in GNSS positioning is generally less precise due to the inherent satellite geometry and atmospheric influences.

Figure 10 illustrates the same analysis as Fig. 9, applied to the obstructed environment (ST2). Compared to ST1, a noticeably higher variability is observed in all components. The horizontal components show standard deviations of approximately 1.6 cm and 2.0 cm for Easting and Northing, respectively, resulting in a combined horizontal dispersion of

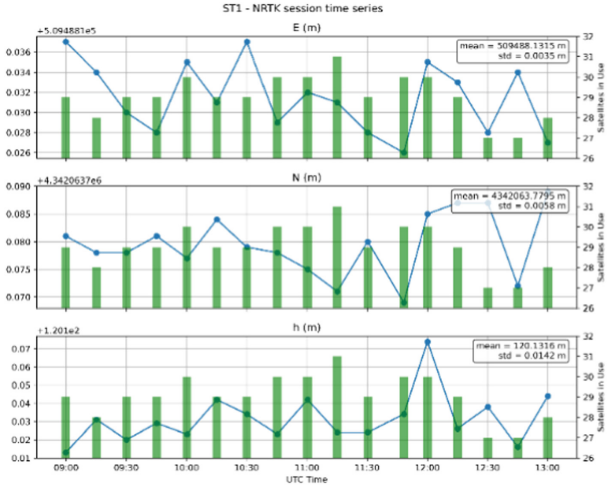


Fig. 9. NRTK time series in the semi-open environment (ST1) for Easting (E), Northing (N), and ellipsoidal height (h), together with the corresponding number of tracked satellites.

about 2.6 cm. This represents a significant degradation with respect to ST1, with the horizontal dispersion increasing by approximately three to four times

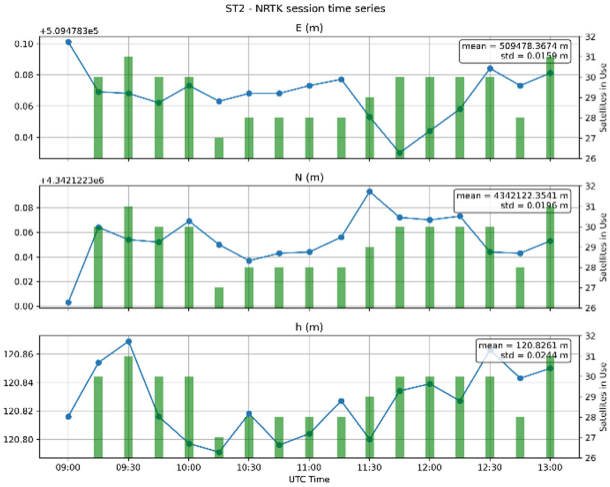


Fig. 10. NRTK time series in the obstructed environment (ST2) for Easting (E), Northing (N), and ellipsoidal height (h), together with the corresponding number of tracked satellites.

The residuals of each epoch with respect to the mean position are generally contained within ± 0.04 m, although several epochs exhibit larger deviations, particularly in the Northing component. This indicates a reduced positioning stability over time, despite all solutions being classified as fixed. The vertical component shows the highest variability,

with a standard deviation of approximately 2.4 cm, confirming the inherently lower precision of the height component in GNSS positioning.

Some noticeable deviations can be identified, especially in the later part of the session (after approximately 12:30 UTC), where both horizontal and vertical components exhibit increased dispersion. These variations are likely related to changes in satellite geometry and to the stronger influence of multipath effects in the obstructed environment, characterized by the presence of buildings and vegetation.

Across both ST1 and ST2, the number of satellites tracked remains relatively stable throughout the observation periods, with only minor variations (Figs. 9 and 10), suggesting no significant correlation between this parameter and the observed performance differences. Instead, the degradation in positioning performance is more plausibly attributed to environmental conditions, such as signal obstructions and multipath effects, as discussed in Sect. 3.2.

To further assess the positioning performance, the NRTK results were compared against the reference solution derived from the complete static session (approximately 4 h) acquired with the GD receiver (see Sect. 2.4). Residuals were computed for the three components as the differences between the NRTK-derived positions and the reference coordinates.

In the semi-open environment (ST1), the NRTK solutions exhibit a good horizontal accuracy, with mean residuals close to zero (−6.0 mm in Easting and − 8 mm in Northing), indicating limited systematic bias. The vertical component, however, exhibits a more pronounced offset, with a mean residual of approximately −6 cm, which dominates the overall 3D error.

In the obstructed environment (ST2), a degradation in performance is observed. The residuals increase in magnitude, especially in the Northing component, reflecting reduced accuracy and enhanced sensitivity to environmental conditions. The vertical component remains the dominant source of error, with a mean bias of approximately − 9 cm (Table 6).

Table 6. Mean residuals of the NRTK solutions with respect to the reference static solution for ST1 and ST2. Values are reported for the Easting (dE), Northing (dN), and ellipsoidal height (dh) components, expressed in meters.

Env	Mean dE (m)	Mean dN (m)	Mean dh (m)
ST1	−0.006	−0.008	−0.06
ST2	0.006	−0.020	−0.09

4 Discussion and Conclusions

The increasing availability of low-cost multi-frequency GNSS receivers has opened new opportunities for high-precision positioning in surveying and geospatial applications. However, their performance under varying environmental conditions, particularly in urban environments, has not yet been fully investigated.

In this context, the performance of a u-blox (LC) GNSS receiver equipped with its patch antenna is analyzed under two environmental conditions representative of typical urban surveying scenarios. Both relative static and Network Real-Time Kinematic (NRTK) positioning strategies are investigated, using positions obtained from long static sessions of a Trimble R12 geodetic device (GD) as a reference.

The analysis of signal quality indicators confirms that the low-cost receiver is more sensitive to environmental conditions compared to the geodetic-grade device. In particular, lower C/N_0 values and higher variability are observed for the LC receiver, especially in the obstructed environment (ST2). This behavior is consistent with the multipath analysis, which shows significantly higher MP values for the LC solution, mainly due to the lower quality of the antenna and its increased susceptibility to signal reflections. The degradation observed from ST1 to ST2 clearly highlights the strong impact of surrounding obstacles on GNSS signal quality. Despite these limitations, the static positioning results indicate that the LC receiver is capable of achieving high accuracy under favorable conditions. In the semi-open environment (ST1), the agreement with the geodetic receiver is at the millimeter level in the horizontal components and at the centimeter level in height. In the obstructed environment (ST2), although fixed solutions are still maintained, the positioning accuracy decreases, with errors reaching the centimeter level, particularly in the Northing and vertical components.

The NRTK analysis further emphasizes the influence of environmental conditions on positioning performance. In ST1, the low-cost receiver provides stable results, with sub-centimeter horizontal dispersion. In contrast, in ST2, a clear degradation is observed, with horizontal dispersion increasing to the centimeter level (approximately 3–4 times higher than in ST1). The vertical component consistently represents the main source of error in both environments, showing systematic biases (approximately -6 cm in ST1 and $-$ 9 cm in ST2) and higher variability.

Our results are fully consistent with the literature on low-cost GNSS receivers, reporting lower C/N_0 values, increased multipath effects, and NRTK performance degradation from millimeter-level accuracy in open-sky conditions to centimeter-level errors in urban environments. Overall, the results demonstrate that low-cost GNSS receivers can be reliably employed in several real-world surveying applications where sub-centimeter performance is not strictly required or where environmental conditions are quite favorable. For instance, these may include topographic and cadastral surveys in open or semi-environments. More recently, such devices have also been proposed for dynamic applications, such as autonomous vehicles and onboard systems for Unmanned Aerial Vehicles (UAV)-based surveying. In these scenarios, operational strategies play a key role in mitigating the observed limitations of low-cost devices, particularly regarding antenna selection and placement.

Future work will focus on a more detailed investigation of LC performance in NRTK mode, including kinematic scenarios, solution stability, ambiguity fixing, and convergence times. Different combinations of receivers and antennas will also be explored, with the aim of identifying optimal configurations for low-cost and high-precision GNSS applications.

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

References

1. Hamza, V., Stopar, B., Sterle, O., et al.: Observations and positioning quality of low-cost GNSS receivers: a review. GPS Solut. **28**, 149 (2024). https://doi.org/10.1007/s10291-024-01686-8
2. Tsakiri, M., Sioulis, A., Piniotis, G.: The use of low-cost, single-frequency GNSS receivers in mapping surveys. Surv. Rev. **50**(358), 46–56 (2018). https://doi.org/10.1080/00396265.2016.1222344
3. Hamza, V., Stopar, B., Sterle, O., et al.: Recent advances and applications of low-cost GNSS receivers: a review. GPS Solut. **29**, 56 (2025). https://doi.org/10.1007/s10291-025-01815-x
4. Sanna, G., Pisanu, T., Garau, S.: Behavior of low-cost receivers in base-rover configuration with geodetic-grade antennas. Sensors. **22**, 2779 (2022). https://doi.org/10.3390/s22072779
5. Poluzzi, L., Tavasci, L., Corsini, F., Barbarella, M., Gandolfi, S.: Low-cost GNSS sensors for monitoring applications. Appl. Geomat. (2020). https://doi.org/10.1007/s12518-019-00268-5
6. Gumilar, I., Sidiq, T., Shihran, R., et al.: Developing and evaluating low-cost GNSS to monitor land subsidence in Bandung Basin. Appl. Geomat. **18**, 15 (2026). https://doi.org/10.1007/s12518-025-00668-w
7. Chen, M., et al.: Assessment of the real-time and rapid precise point positioning performance using geodetic and low-cost GNSS receivers. Remote Sens. **16**(8), 1434 (2024). https://doi.org/10.3390/rs16081434
8. Curone., D., et al.: An innovative low-power, low-cost, multiconstellation geodetic-grade global navigation satellite system reference station for the densification of permanent networks:the GREAT project. Sensors. **23**(13), 6032 (2023). https://doi.org/10.3390/s23136032
9. Hamza, V., Stopar, B., Sterle, O., Pavlovčič-Prešeren, P.: Low-cost dual-frequency GNSS receivers and antennas for surveying in urban areas. Sensors. **23**, 2861 (2023). https://doi.org/10.3390/s23052861
10. u-blox: ZED-F9P module (2026). https://www.u-blox.com/en/product/zed-f9p-module. Last Accessed 20 Mar 2026
11. u-blox: ANN-MB Series GNSS Antennas (2026). https://www.u-blox.com/en/product/ann-mb-series-gnss-antennas. Last Accessed 20 Mar 2026
12. Trimble: Trimble R12 GNSS receiver (2026). https://geospatial.trimble.com/en/products/hardware/trimble-r12. Last Accessed 20 Mar 2026
13. Romero-Andrade, R., Trejo-Soto, M.E., Vega-Ayala, A., et al.: Positioning evaluation of single- and dual-frequency low-cost GNSS receiver signals using PPP and static relative methods in urban areas. Appl. Sci. **11**(22), 10642 (2021)
14. Paziewski, J.: Multi-constellation single-frequency ionospheric-free precise point positioning with low-cost receivers. GPS Solut. (2022). https://doi.org/10.1007/s10291-021-01209-9
15. Ghizzo, E., Lesouple, J., Milner, C.: Assessing spoofing impact on GNSS receivers: carrier-to-noise density ratio (C/N_0) estimation. Signal Process. **241**, 110389 (2026). https://doi.org/10.1016/j.sigpro.2025.110389
16. Jiang, Y., Guo, C., Wang, J., Xu, R.: Multipath effects mitigation in offshore construction platform GNSS-RTK displacement monitoring using parametric temporal convolution network. Remote Sens. **17**(4), 601 (2025). https://doi.org/10.3390/rs17040601
17. Krzan, G., Dawidowicz, K., Paziewski, J.: Low-cost GNSS antennas in precise positioning: a focus on multipath and antenna phase center models. GPS Solut. **28**(3), 103 (2024). https://doi.org/10.1007/s10291-024-01645-3
18. Poluzzi, L., Gandolfi, S.: Performance of dual-frequency low-cost GNSS sensors for real-time monitoring. In: Geomatics and Geospatial Technologies. Communications in Computer and Information Science. Springer (2022). https://doi.org/10.1007/978-3-030-94426-1_18

19. Chen, J., et al.: Performance analysis of a GNSS multipath detection and mitigation method with two low-cost antennas in RTK positioning. IEEE Sensors J. **22**(6), 4827–4835 (2022). https://doi.org/10.1109/JSEN.2021.3068767
20. Hamza, V., Stopar, B., Sterle, O.: Testing the performance of multi-frequency low-cost GNSS receivers and antennas. Sensors. **21**(6), 2021 (2021). https://doi.org/10.3390/s21062029
21. Agenzia Spaziale Italiana (ASI): Nuova rete fiduciale nazionale GNSS dell'Agenzia Spaziale Italiana (2026). https://www.asi.it/2021/11/completata-la-nuova-rete-fiduciale-nazionale-gnss-dellagenzia-spaziale-italiana/. Last Accessed 1 Mar 2026
22. Takasu, T.: RTKLIB: An open source program package for GNSS positioning (2026). http://www.rtklib.com. Last Accessed 1 Jan 2026
23. SARNET: Web server of the Sardinian Permanent GNSS Network (2025). https://www.sarnet.it/servizi.html. Last Accessed 1 Jan 2026



Integration of 3D Ultrasonic Tomography, Photogrammetry, and Mixed Reality for the Structural Assessment of the Marble Statue "Caino"

Marco Vincenzo Majani(✉), Alessandra Carollo, Raffaele Martorana, and Patrizia Capizzi

Università degli Studi di Palermo, Dipartimento di Scienze della Terra e del Mare (DiSTeM), Palermo, Italy
marcovincenzo.majani@unipa.it

Abstract. Increasingly often, museums and art galleries require non-invasive inspections of marble sculptures on the occasion of transfers or museum loans in order to detect any hidden micro-fractures or internal decay. In this regard, we carried out a complete and non-invasive internal assessment using ultrasonic techniques, to ensure the structural stability of the monumental marble statue "Cain" (Domenico Trentacoste, 1903) from the Galleria d'Arte Moderna (GAM) in Palermo, prior to a planned transfer.

The chosen approach integrates high-resolution digital photogrammetry with 3D ultrasonic tomography (UST). Using a series of scripts developed in MATLAB and Python, we reconstructed an ultrasonic wave velocity model strictly constrained to the geometric shape of the artwork. The data acquisition process involved 370 measurement points, including sources and receivers, and among the possible source-receiver pairs, 520 paths were selected to ensure sufficient coverage of all parts of the statue. The results indicated an average pressure wave velocity of approximately 5000 m/s, and the tomographic inversion confirmed an exceptional compactness of the material without any critical internal defects.

Because raw geophysical data can be difficult for conservators to interpret, we exported the 3D velocity model into Mixed and Augmented Reality (MR/AR) environments. This approach allows users to interactively explore cross-sections of the statue in real-time, offering a highly intuitive method for interpreting non-destructive testing data in conservation planning.

Keywords: Cultural Heritage · 3D Ultrasonic Tomography · Photogrammetry · Mixed Reality · Non-Destructive Testing (NDT) · Tomographic Inversion

1 Introduction

Assessing the structural health of monumental works is an important step in preparing for their relocation. In this regard, non-destructive testing (NDT) represents the established standard for examining the internal condition of statues without causing damage [1–4].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 104–114, 2027.
https://doi.org/10.1007/978-3-032-30539-8_7

However, the statue's complex geometry often negatively influences the results of physical models obtained with tomographic methods, which generally require approximate starting models.

Domenico Trentacoste's early twentieth-century masterpiece, "Cain," perfectly illustrates this complexity. Sculpted from fine white statuary marble, the sculpture depicts the biblical figure in a state of profound remorse. His anatomical posture therefore presents protruding volumes and notable morphological variations, causing considerable difficulties in the acquisition of tomographic data and in the subsequent discretization of the model. In view of a planned relocation, the Galleria d'Arte Moderna (GAM) in Palermo requested a comprehensive diagnostic investigation following a preliminary technical inspection conducted in July 2025.

NDT methods struggle to manage particularly complex 3D models, frequently producing distortions in spatial representation and consequent artifacts among the identified anomalies [5, 6]. To minimize these effects, we implemented a diagnostic protocol that combines Structure-from-Motion (SfM) photogrammetry, constrained 3D ultrasonic tomography, and mixed reality (MR) visualization [5, 7, 8]. The resulting model provides a geophysical reconstruction of the elastic wave velocity model, extremely accurate in both form and representation, minimizing artifacts related to geometric distortions. The resulting model represents a very robust tool for assessing the possibility of safely relocating the statue of "Cain," while also providing a powerful tool for visualizing the results for restoration professionals.

2 Related Works

Ultrasound computed tomography (UST) has been a decades-long established technique for assessing the mechanical and physical properties of natural stone [1–3]. The three-dimensional reconstruction of the P-wave velocity field by measuring their propagation times between sources and receivers, placed at appropriate points on the surface, allows for the precise identification of internal defects, ranging from decohesion planes to metallic inclusions [9–12].

However, the highly irregular geometric shape of some sculptures represents a serious obstacle for UST algorithms. Indeed, solving the tomographic inverse problem requires high-precision estimation of the spatial coordinates of the ultrasound sources and receivers [8, 9]. Digital photogrammetry directly addresses this problem, as it allows for the generation of more accurate 3D meshes and, consequently, for the propagation time equations to be constrained to the true physical contours of the artwork, drastically reducing inversion artifacts.

Data interpretation is also evolving. Virtual and mixed reality (VR/MR) platforms are helping to change the way volumetric geophysical data can be represented and viewed. Recent studies in the fields of non-destructive testing (NDT) and geosciences demonstrate that immersive environments recreated by mixed reality offer a much better spatial understanding of 3D information [7, 8, 13]. In the cultural heritage sector, this allows conservators and restorers to intuitively navigate tomographic models [14], significantly improving the decision-making process related to intervention works.

3 Methodology

The methodology discussed in this work is divided into the following phases. First, the measurement points are selected and adhesive markers are used to locate them. Next, a detailed photogrammetric scan is performed to reconstruct a high-detail 3D model of the artwork's surface, including the coordinates of the measurement points. With these spatial constraints, the ultrasonic data is acquired and inverted, using the photogrammetric model as the boundary of the geophysical model. The final phase involves adapting the resulting 3D volume in a VR/MR environment for immersive visualization.

3.1 Photogrammetric Survey

The complex morphology of the "Cain" required careful selection of measurement points to ensure sufficient coverage of the wave paths, which was necessary to obtain an ultrasound velocity model with homogeneous resolution across the entire statue. The digital photogrammetric model of the statue's surface was recreated using RealityScan software [15], which processed 300 high-resolution images, taken from different angles (Fig. 1a), to generate the point cloud (Fig. 1b) together with the textured polygonal mesh (Fig. 1c).

The photogrammetric model obtained and visualized using ParaView [16] was not only for display purposes, but also served a dual purpose in the workflow. The first purpose was to estimate the coordinates of the 370 measurement points (blue dots in Fig. 1c) with extreme precision, ensuring the sensor coordinates coincided with positions on the statue's surface. Furthermore, the generated mesh was used to strictly constrain the subsequent tomographic modeling to the boundary, ensuring that all calculated ray paths remained within the actual volume of the marble, ensuring the physical and geometric consistency of the tomographic inversion process.

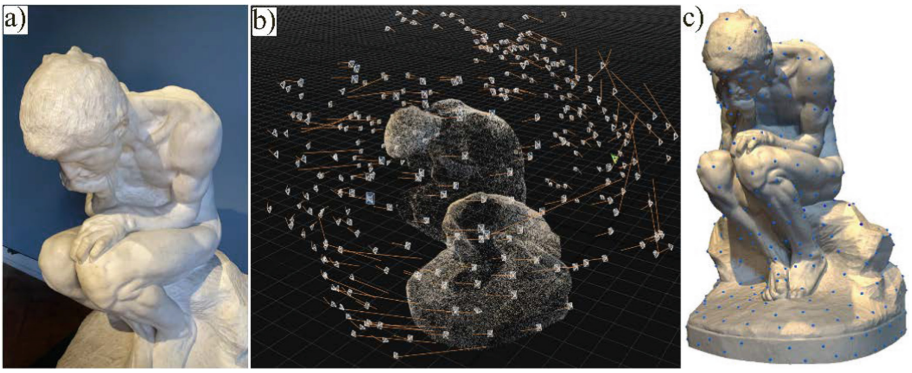


Fig. 1. a) one of the 300 photos of the photogrammetric survey; b) point cloud obtained using RealityScan; c) final textured polygonal mesh of the Caino statue visualized by ParaView. The blue dots indicate the ultrasonic measurement points.

3.2 Ultrasonic Data Acquisition and Tomographic Inversion

3D UST data acquisition was performed using the direct transmission technique. The equipment consisted of a TDAS 16 multichannel system (Boviar S.r.l.), connected to a TSG-55 piezoelectric transmitter and RSG-55 receivers operating at a central frequency of 55 kHz (Fig. 2a). This frequency offers a good compromise between the penetration depth of the ultrasonic pulse and the resolution of the velocity model. The sensors were equipped with conical steel tips specifically designed to ensure localized contact on the curved surface of the marble. However, for optimal acoustic coupling between the sensors and the marble surface, the use of contact gel was necessary to ensure good intensity of the recorded signal. The prohibition of direct contact between the gel and the porous surface of the marble, which could have caused alteration of the latter, was circumvented by applying an ultra-thin, acoustically transparent polyvinyl chloride (PVC) film over the entire surface of the statue (Fig. 2b). The reference markers, necessary to define the acquisition points, were then glued onto the PVC film (Fig. 2c). Finally, ultrasound measurements were taken on the wrapped statue in direct transmission mode (Fig. 2d).

On each acquired ultrasonic signal, the first arrival times of the P waves were manually picked. Any outliers in the dataset were searched for and eliminated by analyzing apparent velocities from the simple ratio between source-receiver distances and the corresponding travel times. A linear regression of travel time versus path length was applied to the filtered dataset to estimate the average apparent velocity (with the straight-ray approximation) to be used as the value for the initial model cells (Fig. 3). The inversion of the estimated travel times was performed using the SIRT (Simultaneous Iterative Reconstruction Technique) algorithm implemented in the GeoTomCG software [17]. This algorithm uses a curvilinear raytracing that takes into account the refractions caused by the velocity variations between the cells into which the model is divided. The solution of the inverse problem starts from an initial model in which each cell of the discretized model is assigned the same average velocity value, obtained from an average of the apparent velocities obtained from the ratio between the travel times and the straight-line distances (homogeneous model and rectilinear raytracing). The model is iteratively updated by the algorithm that tends to minimize the discrepancy between the measured and calculated travel times until a model with minimum stable misfit is obtained. In this research, the optimal model was obtained after seven iterations of the inversion algorithm. A notable feature of this class of inversion methods is that the reconstructed velocity model is generated using a regular Cartesian grid. Although this representation is computationally efficient, it becomes problematic when the actual volume investigated has a very irregular shape, as in most statues. In these cases, the inversion grid necessarily includes regions outside the marble volume, and the corresponding velocity values cannot be interpreted as real material properties. The presence of these external regions can also lead to low-velocity anomalies in the model regions close to the statue's surface, thus complicating structural interpretation if no additional geometric filter is applied. The sequence of steps involved in tomographic inversion is illustrated in the upper section of the flowchart in Fig. 4.



Fig. 2. Phases of ultrasonic data acquisition on the "Caino" statue. a) TDSA acquisition system; b) application of the ultra-thin transparent PVC protective film on the marble surface; c) positioning of reference markers on the wrapped surface to define the measurement points; d) execution of ultrasonic measurements using conical-tip sensors.

3.3 Matlab-Based Data Processing and Geometric Constraint

The most critical phase of the MATLAB routine occurred after the GeoTomCG inversion. The software outputs the reconstructed velocity field as a standard ASCII file, mapping coordinates and velocities across a rigid Cartesian grid. To fix the boundary issues mentioned earlier, we used the high-resolution photogrammetric mesh to physically crop the dataset, stripping away any grid nodes that fell outside the actual marble volume.

Following the SIRT inversion, the resulting ASCII dataset, including spatial coordinates and acoustic velocity values (x, y, z, V), was processed with MATLAB [18], in order to accurately correlate the tomographic data with the boundaries of the statue, obtained from the high-resolution photogrammetric mesh (in OBJ format). Starting from

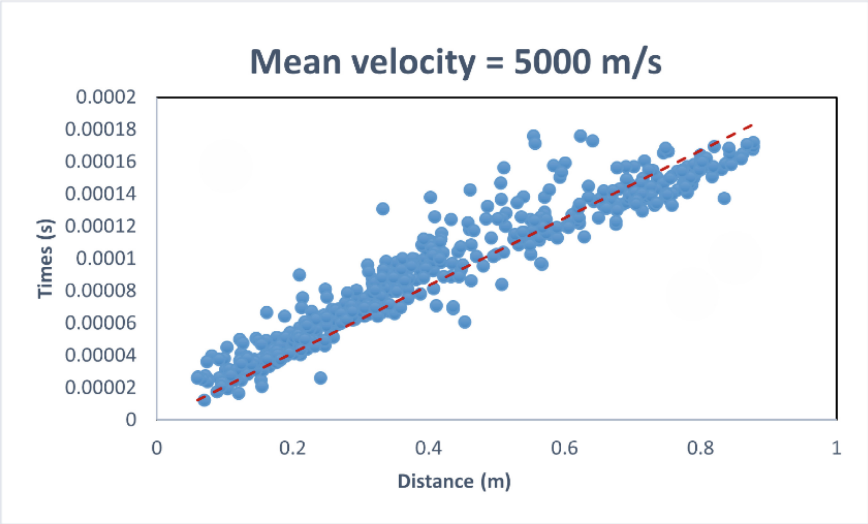


Fig. 3. Travel-time versus source-receiver distance for the filtered first-arrival picks. Linear regression (y = 0.0002x) generate a mean P-wave velocity of 5000 m/s.

the approximate geometry of the inverse model, a regular 3D grid is generated, characterized by a finer spatial resolution (5 mm), whose velocity values are interpolated starting from the velocities of the dataset obtained from the inversion. For the interpolation, the Inverse Distance Weighting (IDW) method was used, through a KD-tree spatial partitioning algorithm, configured with a weighting power equal to 2 and a search radius r of 0.5 m. To strictly limit the tomographic results to the real geometry of the artifact, a geometric constraint is applied using a point-in-mesh test based on a ray-casting algorithm (specifically, the inpolyhedron function). This algorithm evaluates the spatial position of each node in the interpolated grid with respect to the photogrammetric mesh. The voxels of the grid located outside the mesh boundaries are discarded, while the internal voxels are selectively retained. This results in a 3D velocity model completely bounded by the actual geometry of the statue.

The procedure reduced the influence of boundary-related artefacts and improved the physical interpretability of tomographic reconstruction, particularly in geometrically complex areas such as overhangs, curved surfaces, and strongly varying cross-sections. The filtered dataset was then used for three-dimensional rendering together with the external photogrammetric model. The flowchart of the above explained workflow is showed in Fig. 4.

3.4 Augmented and Mixed Reality Representation

To improve the interpretation of the tomographic velocity model, it was integrated into a Mixed Reality (MR) environment. This integration allowed the statue to be visualized at a 1:1 scale, providing the user with a more realistic perception of the spatial distribution of the velocity field within the statue. The 3D velocity model is converted to be usable

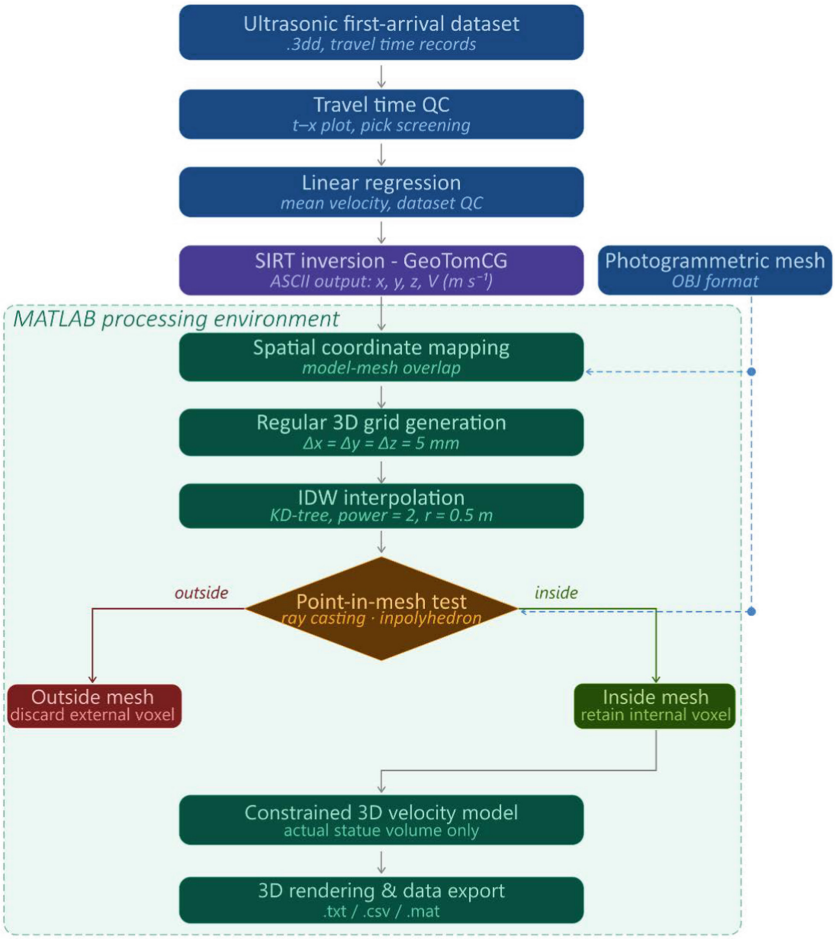


Fig. 4. Algorithmic flowchart of the integrated workflow for geometry-constrained 3D ultrasonic tomography, from first-arrival data preprocessing and SIRT inversion in GeoTomCG to MATLAB-based spatial mapping, IDW interpolation, and point-in-mesh masking of the reconstructed velocity volume using the photogrammetric mesh as geometric constraint.

by modern rendering engines (e.g., Unreal Engine), enabling the user operating in a MR environment to use hand tracking or spatial controllers. This allows for not only static visualizations but also interactive slicing, i.e., the ability to dynamically define, manipulate, and reposition appropriately oriented cross-sections. This can make the analysis of internal velocity anomalies more intuitive and therefore more effective, and consequently, the recognition of any structural inhomogeneities. Ultimately, the use of mixed reality, in addition to being a valid communication and dissemination tool, proves to be an effective support system for diagnosing the state of conservation of the artwork and for subsequent restoration operations.

4 Results and Discussion

The closed triangular mesh produced by the photogrammetric survey rendered the complex anatomical geometry and surface texture of the marble with a precision well above that required both for the georeferencing of the 370 measurement points and for the subsequent refinement of the mesh for the representation of the inverse model in mixed reality.

Preliminary kinematic analysis of the dataset, through the graphical representation of travel times as a function of the corresponding source-receiver distances for all 520 paths, highlighted very few outliers, demonstrating the initial distributional homogeneity of the material. Linear regression of these data showed a global mean P-wave velocity of approximately 5000 m/s, with low variance in the distribution of the data cloud. In the context of statuary marble, this high value indicates a stone matrix characterized by excellent mechanical compactness and a broad structural homogeneity [6].

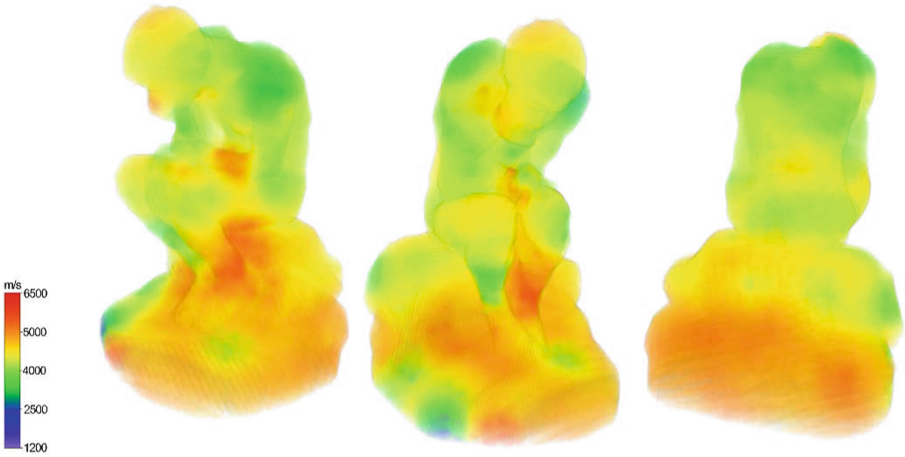


Fig. 5. 3D renderings of the final tomographic velocity model of the Caino statue from three different perspectives. The predominantly green-to-yellow velocity pattern suggests a generally homogeneous and mechanically sound marble matrix, with no evidence of major low-velocity anomalies attributable to significant internal fracturing or voids.

The GeoTomCG inversion produced a satisfactory final misfit between the measured and calculated travel times, considering the low noise level of the dataset, but also taking into account the complexity of the statue's geometry. After reinterpolation of the model mesh by constraining it to the photogrammetric mesh, the resulting 3D velocity model showed a predominantly homogeneous internal distribution (Fig. 5). Most of the reconstructed volume is characterized by velocities exceeding 3000 m/s, consistent with the hypothesis of a compact marble free of widespread fracturing or significant internal decohesion. Volumetric analysis revealed an almost complete absence of anomalies at velocities below 1500–2000 m/s, which would typically indicate severe degradation, hidden cavities, or through fractures.

The transposition of the NDT geophysical data into an Augmented/Mixed Reality environment (Fig. 6) allowed us to interactively obtain cross-sections of the statue's interior, with coordinates and orientation customizable by the user in order to selectively explore the different parts of the statue and highlight any anomalous areas, visually superimposing them on the visual reality.

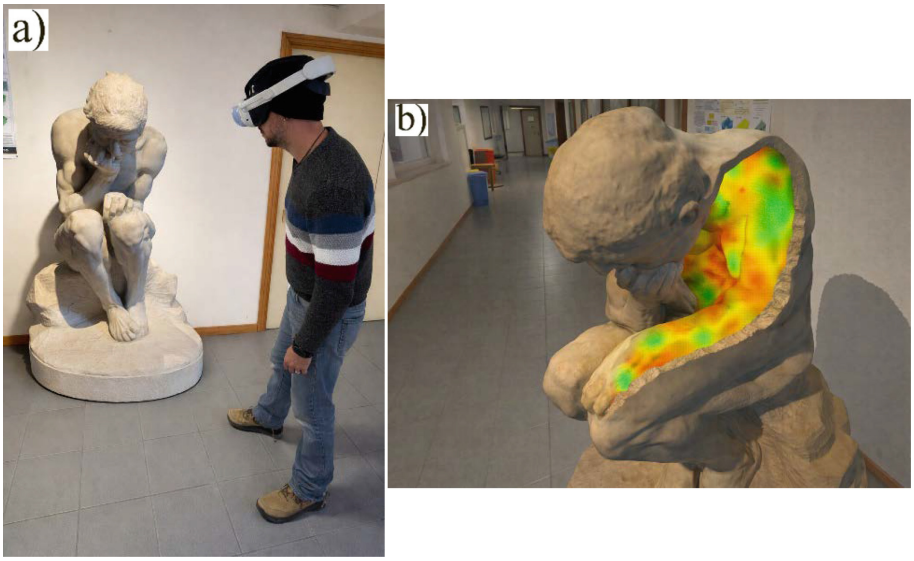


Fig. 6. Mixed Reality visualization of the tomographic model. a) third-person view shows the user observing the statue through the MR headset. b) the corresponding first-person view enables real-time inspection of the statue's internal anomalies, directly superimposed within the 3D model.

5 Conclusions

The present study shows that ultrasonic tomography can provide useful structural information on complex marble sculptures when the tomographic reconstruction is interpreted under an explicit geometric constraint. In the investigated case, the combination of ultrasonic travel-time inversion and photogrammetric modelling made it possible to obtain a three-dimensional velocity distribution that is more consistent with the true volume of the artefact and therefore more reliable for diagnostic interpretation. The main methodological point emerging from the study is not simply the use of photogrammetry alongside tomography, but the use of the photogrammetric mesh as a geometric reference for filtering the regular-grid inversion output. This step is particularly relevant for sculptures and other heritage objects with irregular morphology, where the computational grid extends beyond the actual material boundaries and may introduce non-physical features near the model edges. By removing the external portion of the reconstructed domain, the final model becomes easier to interpret and less exposed to boundary-related artefacts. For the statue *Caino*, the constrained velocity model indicates an overall condition of

internal continuity. The mean ultrasonic velocity is close to 5000 m/s, most of the reconstructed volume shows values above 3000 m/s, and no significant anomalies attributable to through-going fractures, internal cavities, or major degradation zones were detected. Within the limits of the acquisition geometry and tomographic resolution, these results support an interpretation of general structural soundness. These findings do not replace the need for appropriate conservation, transport, and handling protocols. They do, however, provide an objective internal diagnostic basis for evaluating the structural condition of the sculpture in view of possible movement. More generally, the procedure presented here may be applied to other stone artworks whenever tomographic inversion is performed on a regular grid but the investigated object has a non-regular geometry. Finally, integrating the tomographic model into a Mixed Reality framework allowed for a 1:1 scale visualization of the artwork, yielding a significantly enhanced perception of both the artifact's physical proportions and the 3D spatial distribution of the diagnostic results.

Acknowledgments. The authors wish to express their sincere gratitude to the Directorate of the Galleria d'Arte Moderna (GAM) in Palermo for their logistical support and collaboration. This research was conducted within the framework of the Third Mission and Public Engagement activities of the DiSTeM Department, University of Palermo.

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

References

1. Martinho, E., Dionísio, A.: Main geophysical techniques used for non-destructive evaluation in cultural built heritage: a review. J. Geophys. Eng. **11**(5), 053001 (2014). https://doi.org/10.1088/1742-2132/11/5/053001
2. Leucci, G.: Nondestructive testing technologies for cultural heritage: overview. In: Nondestructive Testing for Archaeology and Cultural Heritage, pp. 15–73. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-01899-3_3
3. Martorana, R., Capizzi, P., Pisciotta, A., Scudero, S., Bottari, C.: An overview of geophysical techniques and their potential suitability for archaeological studies. Heritage. **6**(3), 2886–2927 (2023). https://doi.org/10.3390/heritage6030154
4. Piroddi, L. et al.: Imaging cultural heritage at different scales part I, the micro-scale (manufacts). Remote Sens. **15**(10), 2586 (2023). https://doi.org/10.3390/rs15102586
5. Choi, H., Ham, Y., Popovics, J.S.: Integrated visualization for reinforced concrete using ultrasonic tomography and image-based 3-D reconstruction. Constr. Build. Mater. **123**, 384–393 (2016). https://doi.org/10.1016/j.conbuildmat.2016.07.010
6. Capizzi, P., Martorana, R., Carollo, A.: Investigating the internal deterioration of the Auriga statue of Mozia Island, Sicily, through ultrasonic and ground-penetrating radar studies. Sensors. **24**(19), 6450 (2024). https://doi.org/10.3390/s24196450
7. Rehbein, J., Lorenz, S.J., Holtmannspötter, J., Valeske, B.: 3D-visualization of ultrasonic NDT data using mixed reality. J. Nondestruct. Eval. **41**(1), 26 (2022). https://doi.org/10.1007/s10921-022-00860-7
8. Janeras, M. et al.: Using mixed reality for the visualization and dissemination of complex 3D models in geosciences-application to the Montserrat massif (Spain). Geosciences. **12**(10), 370 (2022). https://doi.org/10.3390/geosciences12100370

9. Peterson, J.E., Paulson, B.N.P., McEvilly, T.V.: Applications of algebraic reconstruction techniques to crosshole seismic data. Geophysics. **50**(10), 1566–1580 (1985). https://doi.org/10.1190/1.1441847
10. Phillips, W.S., Fehler, M.C.: Traveltime tomography: a comparison of popular methods. Geophysics. **56**(10), 1639–1649 (1991). https://doi.org/10.1190/1.1442974
11. Jackson, M.J., Tweeton, D.R.: 3DTOM: Three-Dimensional Geophysical Tomography. Report of Investigations 9617, U.S. Bureau of Mines (1996)
12. Lehmann, B.: Seismic Traveltime Tomography for Engineering and Exploration Applications. EAGE Publications, Houton the Netherlands (2007)
13. Alene, G.H. et al.: Virtual reality visualization of geophysical flows: a framework. Environ. Model Softw. **177**, 106063 (2024). https://doi.org/10.1016/j.envsoft.2024.106063
14. Cecotti, H.: Cultural heritage in fully immersive virtual reality. Virtual Worlds. **1**(1), 82–102 (2022). https://doi.org/10.3390/virtualworlds1010006
15. Epic Games.: RealityScan (v.1.8.1/244) [App mobile]. www.capturingreality.com (2024).
16. Sandia National Labs, Kitware Inc., Los Alamos National Labs. ParaView: Parallel Visualization Application. http://paraview.org, 2000–2026 (v.6.0.1). (2026)
17. Dines, K.A., Lytle, R.J.: Computerized geophysical tomography. Proc. IEEE. **67**(7), 1065–1073 (1979). https://doi.org/10.1109/PROC.1979.11390
18. MathWorks Inc.: MATLAB version 24.2. Natick, Massachusetts, United States, (R2024b). (2024). https://www.mathworks.com. Last Accessed 2024

PHD Showcase Papers

An Efficient and Stable Solver for the Shallow Water Equations on Regular Latitude-Longitude Grids Using Spherical Fourier Transforms

Suzune Nomura[1(✉)] and Takeshi Enomoto[2]

[1] Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan
nomura@dpac.dpri.kyoto-u.ac.jp
[2] Disaster Prevention Research Institute, Kyoto University, Uji 611-0011, Japan
enomoto.takeshi.3n@kyoto-u.ac.jp

Abstract. In this study, we apply a spatial differentiation method using the one-dimensional Fourier transform, along with an isotropic wavenumber truncation method based on the direct conversion between spherical Fourier and spherical harmonic coefficients, to the shallow water equations on a sphere. By utilizing the geometric properties of vectors at the poles and evaluating the limit as the colatitude approaches the poles, we successfully resolved the pole singularity problem. Consequently, this approach is applicable to regular latitude-longitude grids that include the poles.

Spatial differentiation is performed in an intermediate space between the physical and wavenumber spaces, without requiring a full transformation into the spherical Fourier space, whereas the wavenumber truncation is executed exclusively within the spherical Fourier space. The computational costs for these procedures are $O(N^2 \log N)$ and $O(N^3)$, respectively. These computations are highly parallelizable, enabling fast computation, and can be implemented without altering the basis functions of the Fourier transform. Furthermore, the intrinsic noise of the spherical Fourier transform is effectively removed by isotropic truncation, ensuring stable time integration.

Numerical experiments applying these methods to the shallow water equations demonstrate that the proposed approach yields smaller errors than conventional spherical harmonic transform methods in both linear and nonlinear problems. Moreover, parallelization using Fortran Coarrays achieved computational speeds comparable to those of the spherical harmonic transform. These findings indicate that our method, which does not require the spherical harmonic transform, is an efficient and highly accurate approach for solving time-dipendent partial differential equations on a sphere, including the polar regions.

Keywords: Pole singularity problem · Isotropic wavenumber truncation · Legendre transformation · Spectral methods on the sphere · Dynamical core · Parallel computing

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 117–134, 2027.
https://doi.org/10.1007/978-3-032-30539-8_8

1 Introduction

In recent years, global atmospheric models worldwide have undergone significant advancements toward higher resolutions. For example, the dynamical core of the Global Spectral Model at the Japan Meteorological Agency was upgraded from TL349L40 to TL959L60 in 2007, and TQ959L128 is currently in operation as of March 2026 (where TL and TQ denote a linear grid and a quadratic grid, respectively). Furthermore, the spectral method is also employed in the models of the European Centre for Medium-Range Weather Forecasts. Because hydrostatic models primarily target large-scale phenomena such as jet streams, the spectral method, which can analytically calculate the derivatives of waves, is an extremely effective approach.

In conventional spectral methods on a sphere, the spherical harmonic transform has been predominantly used. However, its computational cost reaches $O(N^3)$, where N is the truncation wavenumber, making the enormous computational burden a bottleneck for achieving higher resolutions. To avoid this computational cost, an increasing number of global models are transitioning to local discretization schemes such as the finite volume or finite element method. Meanwhile, as an approach to improve computational efficiency while maintaining the high accuracy of wave resolution inherent in spectral methods, the spherical Fourier transform has been proposed [2–4,14,22]. These computational methods have a significant advantage in that their computational cost is $O(N^2 \log N)$ and they can utilize the widely used fast Fourier transform (FFT; [5,19]).

There have been several reports of constructing shallow water equation models using spherical Fourier transforms [17,24]. However, the method proposed by Spotz et al. [17] multiplied latitude-dependent metric coefficients (such as $1/\sin\theta$, where θ is the colatitude) on the physical grid. These coefficients appear in longitudinal differentiation on a sphere and cause singularities at the poles, making it impossible to place grid points exactly at the poles. Additionally, since Spotz et al. [17] relied on the spherical harmonic transform to achieve high accuracy equivalent to conventional spectral methods, the inherent computational bottleneck of the transform remained fundamentally unresolved. Unique wavenumber truncations based on the spherical Fourier transform were also proposed; however, these suffered from inferior computational accuracy and the inability to maintain isotropy on the sphere. In contrast, Yoshimura [24] introduced basis functions that include the latitudinally dependent coefficients and implemented them using the Galerkin method, successfully resolving the issue of being unable to place grid points at the poles. Nevertheless, the problem of isotropy in wavenumber truncation remains unresolved.

To address these remaining issues, in this paper, we derive a new formulation to accurately evaluate the gradient, vorticity, and divergence at the poles. Furthermore, for the isotropic wavenumber truncation method, we use a direct conversion from spherical Fourier coefficients to spherical harmonic coefficients, and apply a triangular truncation that mathematically guarantees isotropy on the sphere. Through these advancements, we propose a new highly accurate and

fast solver on a grid including the poles, constructed solely with Fourier transforms and coefficient conversions, thereby achieving computational efficiency.

The structure of this paper is as follows. Section 2 describes the formulation of spatial differentiation at the poles. Section 3 details the wavenumber truncation method that guarantees isotropy on the sphere. In Sect. 4, the new pseudospectral method based on the spherical Fourier transform, which combines the spatial differentiation method from Sect. 2 and the wavenumber truncation from Sect. 3, is applied to the shallow water equations. We then verify the accuracy of the proposed method through comparisons with the conventional spectral method using the spherical harmonic transform, as well as with a pseudospectral method that uses the spherical Fourier transform for spatial differentiation but the spherical harmonic transform for wavenumber truncation. Section 5 compares the execution times between the new spherical Fourier transform and the spherical harmonic transform, using a benchmark test that simplifies the computational process of the shallow water equations. Finally, Sect. 6 presents the conclusions.

2 Formulation on the Poles

For spatial differentiation away from the poles, similar to the method of Spotz et al. [17], one-dimensional Fourier transforms are applied along lines of latitude or longitude, differentiation operations are performed in wavenumber space, and the results are inversely transformed to obtain the derivative values in physical space. In this process, coefficients dependent on the radius of the latitude circle, which arise in calculations such as the longitudinal component of the gradient, are multiplied in physical space.

For differentiation in the longitudinal direction, periodicity is inherently guaranteed; therefore, the one-dimensional Fourier transform can be applied directly. For differentiation in the latitudinal direction, on the other hand, the latitudinal data array at a given longitude λ and the array at longitude $\lambda + \pi$ (shifted by π rad) are concatenated to form a periodic great circle passing through the North and South Poles. This makes it possible to apply the one-dimensional Fourier transform.

The method of Spotz et al. [17] targeted a regular latitude-longitude grid that does not include the poles, and prioritized ensuring a unified direction of differentiation along the great circle passing through the two poles. Consequently, when concatenating arrays across the poles, the signs for scalar fields were inverted, whereas those for vector fields were not. However, since our proposed method is designed to accommodate a grid that includes the poles, it is strictly required that the physical quantities be mathematically continuous at the poles to prevent the Gibbs phenomenon at these singular points. Due to the geometric properties of the spherical coordinate system, when a vector component (latitudinal or longitudinal wind velocity) crosses the pole to a longitude shifted by π rad, the orientation of the local coordinate axis is reversed, and the signs of concatenated vectors are also reversed. Therefore, to preserve physical

continuity at the poles, it is necessary to invert the signs of the concatenated arrays across the poles. This consideration is equivalent to reversing the signs of the derivatives in regions where the longitude is shifted by π, We employed this method.

2.1 The Longitudinal Component of the Gradient at the Poles

Denote the unit vector in Cartesian coordinates by ($\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$) and the unit vector in spherical coordinates by ($\mathbf{e}_r$, $\mathbf{e}_\theta$, $\mathbf{e}_\lambda$); this transformation can be expressed as follows:

$$\begin{pmatrix} \mathbf{e}_r \\ \mathbf{e}_\theta \\ \mathbf{e}_\lambda \end{pmatrix} = \mathbf{R} \begin{pmatrix} \mathbf{i} \\ \mathbf{j} \\ \mathbf{k} \end{pmatrix} = \mathbf{R}_y(\theta)\mathbf{R}_z(\lambda) \begin{pmatrix} \mathbf{i} \\ \mathbf{j} \\ \mathbf{k} \end{pmatrix} \tag{1}$$

where $\mathbf{R}_y(\theta)$ and $\mathbf{R}_z(\lambda)$ are the rotation matrices for the y-axis and z-axis, respectively. The transformation matrix $\mathbf{R}$ is given by

$$\mathbf{R} = \mathbf{R}_y(\theta)\mathbf{R}_z(\lambda) = \begin{pmatrix} \sin\theta & 0 & \cos\theta \\ \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \cos\lambda & \sin\lambda & 0 \\ -\sin\lambda & \cos\lambda & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

In particular, the transformation at the poles is given by

$$\begin{pmatrix} \mathbf{e}_r \\ \mathbf{e}_\theta \\ \mathbf{e}_\lambda \end{pmatrix} = \begin{cases} \mathbf{R}_y(0)\mathbf{R}_z(\lambda) \begin{pmatrix} \mathbf{i} \\ \mathbf{j} \\ \mathbf{k} \end{pmatrix} & \text{as } \theta \to 0 \text{ (north pole)} \\ \mathbf{R}_y(\pi)\mathbf{R}_z(\lambda) \begin{pmatrix} \mathbf{i} \\ \mathbf{j} \\ \mathbf{k} \end{pmatrix} & \text{as } \theta \to \pi \text{ (south pole)} \end{cases}$$

Therefore, at the North Pole, we have

$$\begin{aligned} \begin{pmatrix} \mathbf{e}_r \\ \mathbf{e}_\theta \\ \mathbf{e}_\lambda \end{pmatrix} &= \mathbf{R}_y(0)\mathbf{R}_z(\lambda) \begin{pmatrix} \mathbf{i} \\ \mathbf{j} \\ \mathbf{k} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \cos\lambda & \sin\lambda & 0 \\ -\sin\lambda & \cos\lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{i} \\ \mathbf{j} \\ \mathbf{k} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 1 \\ \cos\lambda & \sin\lambda & 0 \\ -\sin\lambda & \cos\lambda & 0 \end{pmatrix} \begin{pmatrix} \mathbf{i} \\ \mathbf{j} \\ \mathbf{k} \end{pmatrix} \end{aligned}$$

Focusing on $\mathbf{e}_\theta$ and $\mathbf{e}_\lambda$, we obtain

$$\begin{pmatrix} \mathbf{e}_\theta \\ \mathbf{e}_\lambda \end{pmatrix} = \begin{pmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{pmatrix} \begin{pmatrix} \mathbf{i} \\ \mathbf{j} \end{pmatrix}$$

Thus, $\mathbf{e}_\theta$ and $\mathbf{e}_\lambda$ at the North Pole are expressed as

$$\begin{aligned}\mathbf{e}_\theta(0,\,\lambda) &= \cos\lambda\mathbf{i} + \sin\lambda\mathbf{j}\\ \mathbf{e}_\lambda(0,\,\lambda) &= -\sin\lambda\mathbf{i} + \cos\lambda\mathbf{j} = \mathbf{e}_\theta\left(\lambda + \frac{\pi}{2}\right)\end{aligned} \tag{2}$$

Similarly, in South Pole are expressed as

$$\begin{aligned}\mathbf{e}_\theta(\pi,\,\lambda) &= -\cos\lambda\mathbf{i} - \sin\lambda\mathbf{j}\\ \mathbf{e}_\lambda(\pi,\,\lambda) &= -\sin\lambda\mathbf{i} + \cos\lambda\mathbf{j} = -\mathbf{e}_\theta\left(\lambda + \frac{\pi}{2}\right)\end{aligned} \tag{3}$$

Eqs. (2) and (3), considered on the unit circle, is applied to the gradient of an arbitrary scalar $f(\theta, \lambda)$.

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial \theta} \\ \frac{1}{\sin\theta}\frac{\partial f}{\partial \lambda} \end{pmatrix} = \begin{pmatrix} \nabla f \cdot \mathbf{e}_\theta \\ \nabla f \cdot \mathbf{e}_\lambda \end{pmatrix} \tag{4}$$

Substituting Eqs. (2) and (3) into Eq. (4) obtains the longitude component of the gradient at the pole:

$$\begin{aligned}\lim_{\theta\to 0}\frac{1}{\sin\theta}\frac{\partial f}{\partial\lambda} &= \nabla f\cdot\mathbf{e}_\lambda\,(0,\,\lambda) = \nabla f\cdot\mathbf{e}_\theta\left(0,\,\lambda+\frac{\pi}{2}\right)\\ \lim_{\theta\to \pi}\frac{1}{\sin\theta}\frac{\partial f}{\partial\lambda} &= \nabla f\cdot\mathbf{e}_\lambda\,(\pi,\,\lambda) = -\nabla f\cdot\mathbf{e}_\theta\left(\pi,\,\lambda+\frac{\pi}{2}\right)\end{aligned} \tag{5}$$

This means that the longitudinal component of the gradient at the poles is a rotation of the latitudinal component.

2.2 The Vorticity and Divergence at the Poles

The vorticity ζ and divergence δ are fundamentally defined as the limits of circulation C and flux F per unit area ΔS, respectively. By substituting the differential forms of the operators in spherical coordinates and considering the surface element $\mathrm{d}S = \sin\theta\mathrm{d}\theta\mathrm{d}\lambda$ on the unit sphere, these definitions can be expressed as follows:

$$\begin{aligned}\zeta &= \lim_{\Delta S\to 0}\frac{\Delta C}{\Delta S} = \lim_{D\to 0}\frac{\int\int_D \frac{1}{\sin\theta}\left[\frac{\partial(\sin\theta u)}{\partial\theta} - \frac{\partial v}{\partial\lambda}\right]\sin\theta\mathrm{d}\theta\mathrm{d}\lambda}{\int\int_D \sin\theta\mathrm{d}\theta\mathrm{d}\lambda}\\ \delta &= \lim_{\Delta S\to 0}\frac{\Delta F}{\Delta S} = \lim_{D\to 0}\frac{\int\int_D \frac{1}{\sin\theta}\left[\frac{\partial(\sin\theta v)}{\partial\theta} + \frac{\partial u}{\partial\lambda}\right]\sin\theta\mathrm{d}\theta\mathrm{d}\lambda}{\int\int_D \sin\theta\mathrm{d}\theta\mathrm{d}\lambda}\end{aligned} \tag{6}$$

where $\mathbf{v}^{\mathrm{T}} = (u, v)$ is the velocity vector. The cancellation of $\sin\theta$ terms within this integral representation ensures that ζ and δ remain finite, providing a robust

mathematical basis for our treatment of the poles and justifying their global representation through the following limits:

$$\zeta = \begin{cases} \lim\limits_{\theta\to 0} \dfrac{\int_0^{2\pi}\int_0^{\theta} \frac{1}{\sin\theta'}\left[\frac{\partial\left(\sin\theta' u\right)}{\partial\theta'} - \frac{\partial v}{\partial\lambda}\right]\sin\theta' \mathrm{d}\theta'\mathrm{d}\lambda}{\int_0^{2\pi}\int_0^{\theta}\sin\theta'\mathrm{d}\theta'\mathrm{d}\lambda} & \text{for } \theta = 0 \\ \dfrac{1}{\sin\theta}\left[\dfrac{\partial\left(\sin\theta u\right)}{\partial\theta} - \dfrac{\partial v}{\partial\lambda}\right] & \text{for } 0 < \theta < \pi \\ \lim\limits_{\theta\to \pi} \dfrac{\int_0^{2\pi}\int_0^{\theta} \frac{1}{\sin\theta'}\left[\frac{\partial\left(\sin\theta' u\right)}{\partial\theta'} - \frac{\partial v}{\partial\lambda}\right]\sin\theta' \mathrm{d}\theta'\mathrm{d}\lambda}{\int_0^{2\pi}\int_0^{\theta}\sin\theta'\mathrm{d}\theta'\mathrm{d}\lambda} & \text{for } \theta = \pi \end{cases} \tag{7}$$

$$\delta = \begin{cases} \lim\limits_{\theta\to 0} \dfrac{\int_0^{2\pi}\int_0^{\theta} \frac{1}{\sin\theta}\left[\frac{\partial(\sin\theta v)}{\partial\theta} + \frac{\partial u}{\partial\lambda}\right]\sin\theta \mathrm{d}\theta\mathrm{d}\lambda}{\int_0^{2\pi}\int_0^{\theta}\sin\theta\mathrm{d}\theta\mathrm{d}\lambda} & \text{for } \theta = 0 \\ \dfrac{1}{\sin\theta}\left[\dfrac{\partial\left(\sin\theta v\right)}{\partial\theta} + \dfrac{\partial u}{\partial\lambda}\right] & \text{for } 0 < \theta < \pi \\ \lim\limits_{\theta\to \pi} \dfrac{\int_0^{2\pi}\int_0^{\theta} \frac{1}{\sin\theta'}\left[\frac{\partial\left(\sin\theta' v\right)}{\partial\theta'} + \frac{\partial u}{\partial\lambda}\right]\sin\theta' \mathrm{d}\theta'\mathrm{d}\lambda}{\int_0^{2\pi}\int_0^{\theta}\sin\theta'\mathrm{d}\theta'\mathrm{d}\lambda} & \text{for } \theta = \pi \end{cases} \tag{8}$$

Applying L'Hôpital's rule to evaluate this limit yields for the vorticity and divergence at the poles, respectively:

$$\zeta = \begin{cases} \dfrac{1}{\pi}\displaystyle\int_0^{2\pi} \left.\frac{\partial u}{\partial\theta}\right|_{\theta=0} \mathrm{d}\lambda & \text{for } \theta = 0 \\ \dfrac{1}{\pi}\displaystyle\int_0^{2\pi} \left.\frac{\partial u}{\partial\theta}\right|_{\theta=\pi} \mathrm{d}\lambda & \text{for } \theta = \pi \end{cases} \tag{9}$$

$$\delta = \begin{cases} \dfrac{1}{\pi}\displaystyle\int_0^{2\pi} \left.\frac{\partial v}{\partial\theta}\right|_{\theta=0} \mathrm{d}\lambda & \text{for } \theta = 0 \\ \dfrac{1}{\pi}\displaystyle\int_0^{2\pi} \left.\frac{\partial v}{\partial\theta}\right|_{\theta=\pi} \mathrm{d}\lambda & \text{for } \theta = \pi \end{cases} \tag{10}$$

This result is consistent with the fact that, at the poles, the longitudinal component u represents rotation and the latitudinal component v represents divergence, meaning that vorticity and divergence can be expressed solely in terms of these respective components.

2.3 The Laplacians at the Poles

The scalar Laplacian and the vector Laplacian are expressed as follows, respectively [20]:

$$\begin{aligned} \Delta_s f &= \nabla \cdot (\nabla f) \\ \Delta_v \mathbf{f} &= \nabla(\nabla \cdot \mathbf{f}) + \mathbf{k} \times \nabla \left(\mathbf{k} \cdot \nabla \times \mathbf{f}\right) \end{aligned} \tag{11}$$

Here, f and $\mathbf{f}$ are represent an arbitrary scalar and vector, respectively, where $\mathbf{k}$ is the unit vector in the z-axis (i.e., $\mathbf{k}^{\mathrm{T}} = (0, 0, 1)$). While the scalar and vector Laplacians typically involve singularities at the poles, it is possible to obtain finite values by employing formulations for the gradient, vorticity, and divergence that do not diverge (see Sect. 2.1 and 2.2).

3 Formulation of Isotropic Wavenumber Truncation

In previous studies, when solving time-dependent partial differential equations on the sphere using the spherical Fourier transform, pseudo-spectral methods (e.g., [3,17]) and Galerkin spectral methods (e.g., [24]) have been employed. Both approaches require wavenumber truncation to remove high-frequency components generated by nonlinear terms. Wavenumber truncation based on the 3/2 rule in Fourier space for de-aliasing exhibit the following characteristics:

- It cannot entirely eliminate the inherent noise errors near the poles caused by the use of the spherical Fourier transform.
- While spherical harmonics guarantee isotropy on a sphere, Fourier series only guarantee isotropy on a torus.

Due to these characteristics, the importance of applying a filter based on the spherical harmonic transform has been pointed out when using the spherical Fourier transform for spatial differentiation.

3.1 Isotropic Wavenumber Truncation of Scalar Fields

The associated Legendre functions, which constitute spherical harmonics, can be Fourier-transformed [11,12,16]. Consequently, spherical harmonic coefficients can be derived directly from Fourier coefficients [9,10].

Let $\mathrm{ar}_{n,m}$ and $\mathrm{ai}_{n,m}$ be the spherical harmonic coefficients, and $A_{k,m}$ and $B_{k,m}$ be the spherical Fourier coefficients defined by Yee [22]. The transformation between these coefficients is expressed as:

$$A_{k,m} = \sum_{\substack{n=\max(m,k) \\ n-k \in \mathrm{even}}}^{N} p_{n,m,k} \mathrm{ar}_{n,m}, \qquad B_{k,m} = \sum_{\substack{n=\max(m,k) \\ n-k \in \mathrm{even}}}^{N} p_{n,m,k} \mathrm{ai}_{n,m} \tag{12}$$

and the corresponding inverse transform:

$$\mathrm{ar}_{n,m} = \sum_{k=0}^{N} q_{n,m,k} A_{k,m}, \qquad \mathrm{ai}_{n,m} = \sum_{k=0}^{N} q_{n,m,k} B_{k,m} \tag{13}$$

where, $p_{n,m,k}$ and $q_{n,m,k}$ are a function that does not depend on colatitude θ or longitude λ, a precomputable array.

For a triangular truncation with a wavenumber $N_{\mathrm{TR}}(\leq N)$, it suffices to evaluate the transformation up to $n = N_{\mathrm{TR}}$. Therefore, the wavenumber truncation for a given scalar function f is performed according to the following procedure:

1. Apply the spherical Fourier transform to the scalar function $f(\theta, \lambda)$. For a given zonal wavenumber m $(0 \leq m \leq N_{\mathrm{TR}})$, we obtain the coefficients $A_{k,m}$ and $B_{k,m}$ $(0 \leq k \leq N_{\mathrm{TR}})$, which satisfy:

$$\tilde{f}(\theta, \lambda) = \sum_{m=0}^{N_{\mathrm{TR}}} \sum_{k=0}^{N_{\mathrm{TR}}} h_m(k\theta) \left(A_{k,m} \cos m\lambda + B_{k,m} \sin m\lambda\right)$$

   where N_{TR} is the maximum wavenumber and

$$h_m(\psi) \equiv \begin{cases} \cos\psi & \text{if } m \text{ is even} \\ \sin\psi & \text{if } m \text{ is odd} \end{cases}$$

   is a function defined conditionally based on the parity of m.
2. Transform $A_{k,m}$ and $B_{k,m}$ into $\mathrm{ar}_{n,m}$ and $\mathrm{ai}_{n,m}$.
3. Reconstruct $\tilde{A}_{k,m}$ and $\tilde{B}_{k,m}$ from the truncated spherical harmonic coefficients obtained in Step 2.
4. Perform the inverse spherical Fourier transform using $\tilde{A}_{k,m}$ and $\tilde{B}_{k,m}$ obtain the filtered physical data.

The coefficients $\tilde{A}_{k,m}$ and $\tilde{B}_{k,m}$ obtained in Step 3 are generally distinct from the original $A_{k,m}$ and $B_{k,m}$ in Step 2. This is because an associated Legendre function is represented by a finite cosine series, meaning that a single spherical harmonic component is projected onto multiple wavenumbers within the spherical Fourier space. By performing the transformations in Step 2 and Step 3, the resulting coefficients $\tilde{A}_{k,m}$ and $\tilde{B}_{k,m}$ are no longer affected by the projection of spectra that reside in the truncated region of the spherical harmonic space.

3.2 Isotropic Wavenumber Truncation of Vector Fields

By leveraging the isotropic wavenumber truncation for scalar fields described in Sect. 3.1, the truncation of vector fields can be achieved through Helmholtz decomposition. This process is performed according to the following procedure:

1. Compute the vorticity and divergence from the wind velocity vector.
2. Apply the spherical Fourier transform to the vorticity and divergence to obtain their respective spherical Fourier coefficients, $A_{k,m}$ and $B_{k,m}$.
3. Convert $A_{k,m}$ and $B_{k,m}$ into the spherical harmonic coefficients, $\mathrm{ar}_{n,m}$ and $\mathrm{ai}_{n,m}$.
4. Divide $\mathrm{ar}_{n,m}$ and $\mathrm{ai}_{n,m}$ by $-n(n+1)$. This step corresponds to the inverse Laplacian operation in spherical harmonic space.

5. Convert the resulting spherical harmonic coefficients back into the spherical Fourier coefficients, $\tilde{A}_{k,m}$ and $\tilde{B}_{k,m}$.
6. Transform $\tilde{A}_{k,m}$ and $\tilde{B}_{k,m}$ back into physical space using the inverse spherical Fourier transform to obtain the triangularly truncated stream function and velocity potential.
7. Reconstruct the truncated wind velocity vector by calculating the spatial derivatives of the stream function and velocity potential.

The inverse Laplacian operation described above cannot be replaced by solving the Poisson equation directly using the spherical Fourier transform method proposed by Yee [23]. Our numerical experiments confirm that the approach using Yee's Poisson solver induces numerical instability within only a few time steps of integration.

4 Evaluation Using Shallow-Water Test Cases

We conducted Test Cases 1 and 2 proposed by Williamson [21]. Test Case 1 is a linear advection problem, whereas Test Case 2 is a nonlinear steady-state problem. The three methods compared are as follows:

(i) **New spherical Fourier transform method.**
A pseudo-spectral method combining the one-dimensional Fourier transform described in Sect. 2 with the isotropic wavenumber truncation described in Sect. 3.

(ii) **Combined method.**
A pseudo-spectral method combining the one-dimensional Fourier transform described in Sect. 2 with the isotropic wavenumber truncation based on the spherical harmonic transform.

(iii) **Spherical harmonic transform method.**
A Galerkin spectral method based on the spherical harmonic transform.

The third method is one of the standard approaches widely utilized in atmospheric dynamics.

4.1 Governing Equations

The shallow water equations on a sphere can be expressed as follows:

$$\begin{aligned}
\frac{\partial u}{\partial t} &= -u\delta - \frac{v}{a}\frac{\partial u}{\partial \theta} + \frac{u}{a}\frac{\partial v}{\partial \theta} + fv - \frac{1}{a\cos\theta}\frac{\partial \phi}{\partial \lambda} \\
\frac{\partial v}{\partial t} &= -u\zeta - \frac{u}{a}\frac{\partial u}{\partial \theta} - \frac{v}{a}\frac{\partial v}{\partial \theta} - fu - \frac{1}{a}\frac{\partial \phi}{\partial \theta} \\
\frac{\partial \phi}{\partial t} &= -\phi\delta - \frac{u}{a\cos\theta}\frac{\partial \phi}{\partial \lambda} - \frac{v}{a}\frac{\partial \phi}{\partial \theta}
\end{aligned} \tag{14}$$

where t is time, u and v are the zonal and meridional wind components, and ϕ is the geopotential height. The terms ζ and δ denote vorticity and divergence,

respectively. The constants g, a, and f represent the acceleration due to gravity, the radius of the Earth, and the Coriolis parameter. In particular, the geopotential height is defined as:

$$\phi = gh^\star = g(h - h_s) \tag{15}$$

where $h^\star$ is the depth of the fluid and h is the height of the free surface above a reference sphere. All methods employed a regular latitude-longitude grid. The numerical libraries utilized were FFTW (v3.3.10; [8]) and SPHEREPACK 3.2 [1]. We used SPHEREPACK 3.2 because it includes the spherical harmonic expansion method on a regular latitude-longitude grid proposed by Swarztrauber [18].

4.2 Williamson Test Case 1

Test Case 1 is to simulate cosine bell advection. The cosine bell was initially positioned on the equator at $\lambda = \pi/2$. The radius and height of the bell were set to $1/3$ rad and 1000 m, respectively. The bell is advected around the sphere, completing one revolution in 12 days while crossing both poles. Only the geopotential height is evolved over time, without evolving the horizontal wind components. The initial state is shown in Fig. 1.

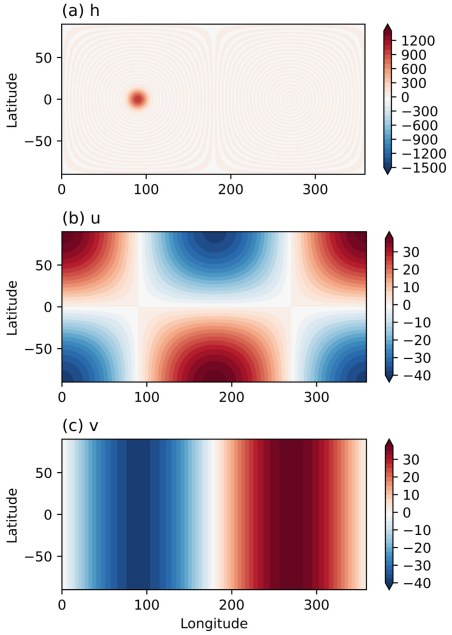


Fig. 1. Initial states for Test Case 1 using the proposed method with a truncation wavenumber of 85. (a) geopotential height, (b) zonal and (c) meridional wind components.

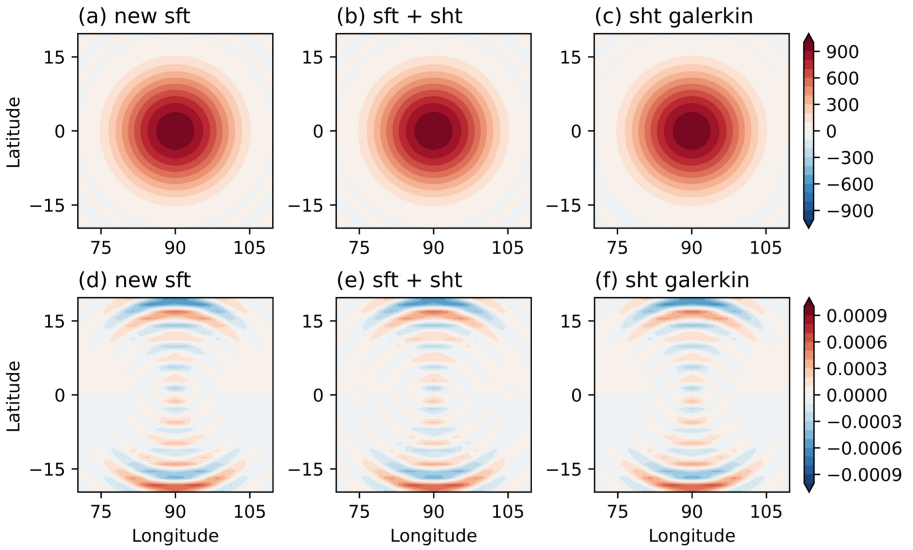


Fig. 2. Simulated states after 12 days for Test Case 1. The truncation wavenumber is 85. (a,d) the proposed method, (b,e) the combined method and (c,f) the spherical harmonic method, and (a–c) and (d–f) show the geopotential height and the error distribution of the geopotential height, respectively.

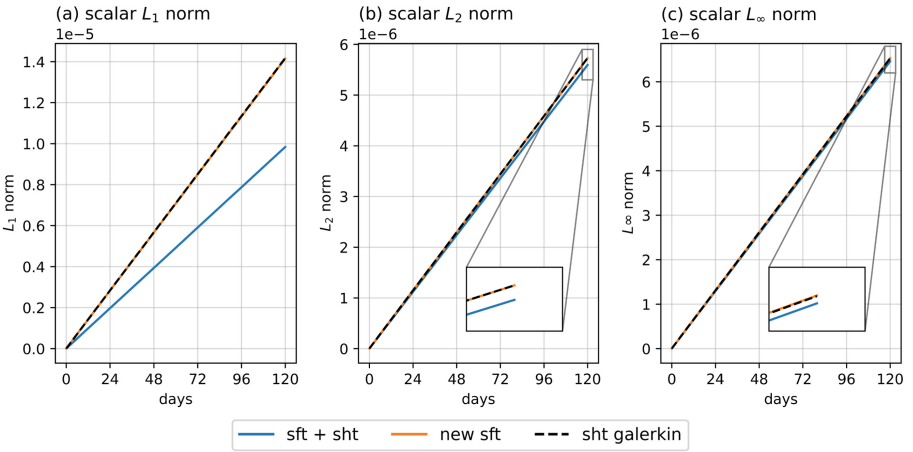


Fig. 3. Error norms for Test Case 1 over a 120 day integration. The truncation wavenumber is 85.

While error patterns show little difference among the methods after 12 days (Fig. 2), the proposed method demonstrates superior accuracy in the error norms over a 120-day integration (Fig. 3). These scalar error norms are evaluated fol-

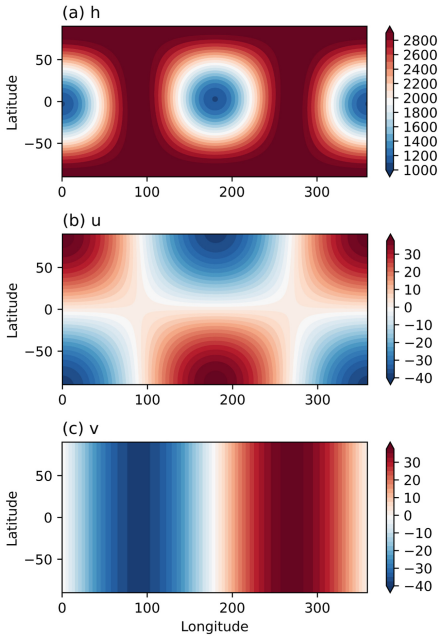


Fig. 4. Same as Fig. 1 but for Test Case 2.

lowing Eqs. (82)–(84) of [21]. Note that the initial conditions for all methods were pre-truncated in wavenumber space.

4.3 Williamson Test Case 2

Test Case 2 is a nonlinear advection experiment that maintains an initial state in geostrophic balance. In this experiment, the angle between the flow's axis of rotation and the Earth's rotation axis was set to $\pi/2 - 0.05$ to intentionally introduce a discontinuity in the wind vector at the poles (Fig. 4).

Regarding the horizontal error distribution of the geopotential height, the proposed method and the combined method show very similar patterns (Fig. 5a,c). In contrast, the spherical harmonic method exhibits prominent noise parallel to the latitude lines (Fig. 5e). For the error distribution of the horizontal wind component, on the other hand, the combined method and the spherical harmonic method show similar characteristics, whereas the proposed method exhibits an error pattern characterized by the zonal wavenumber 1 mode (Fig. 5b,d,f, v is similar to u in behavior).One possible cause of this error pattern is that the vector wavenumber truncation procedure (Sect. 3.2) requires multiple intermediate transformations between the vorticity-divergence fields and the velocity components. These additional steps, compared to the simpler scalar truncation, make the vector field more susceptible to the accumulation of numerical rounding errors.

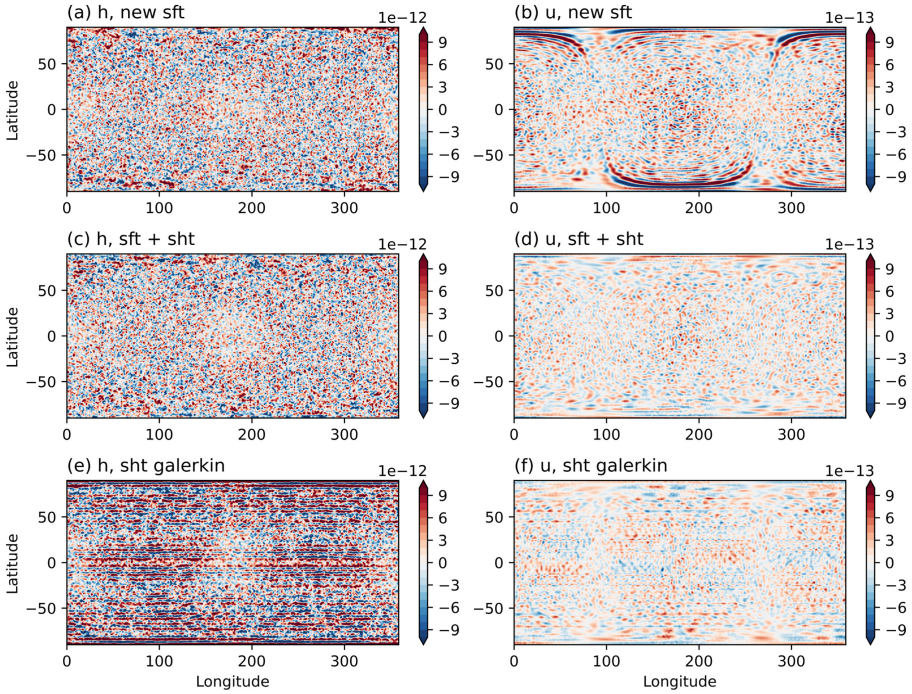


Fig. 5. Error distribution after 10 days for Test Case 2. The truncation wavenumber is 85. (a,b) the proposed method, (c,d) the combined method, and (e,f) the spherical harmonic method. Panels (a,c,e) and (b,d,f) show the error distribution of the geopotential height and zonal wind, respectively.

Nevertheless, evaluating both scalar and vector error norms using Eqs. (97)–(99) of [21] reveals that the proposed method maintains high accuracy in scalar fields even if the vector precision is lower than that of other methods. This demonstrates that the overall utility of the proposed method remains robust (Fig. 6).

5 Preliminary Benchmark of Computation Time

The proposed method consists of the spatial differentiation method described in Sect. 2 and the wavenumber truncation method described in Sect. 3. The spatial differentiation can be computed independently along latitudes and longitudes. Similarly, the wavenumber truncation can derive spherical Fourier coefficients independently in the latitudinal and longitudinal directions. Furthermore, the mutual transformation with spherical harmonic coefficients can be calculated independently for each zonal and meridional wavenumber, resulting in highly efficient parallelization. To demonstrate this advantage, we compared our method with SHTns [15], which implements the world's fastest spherical harmonic transform algorithm [13].

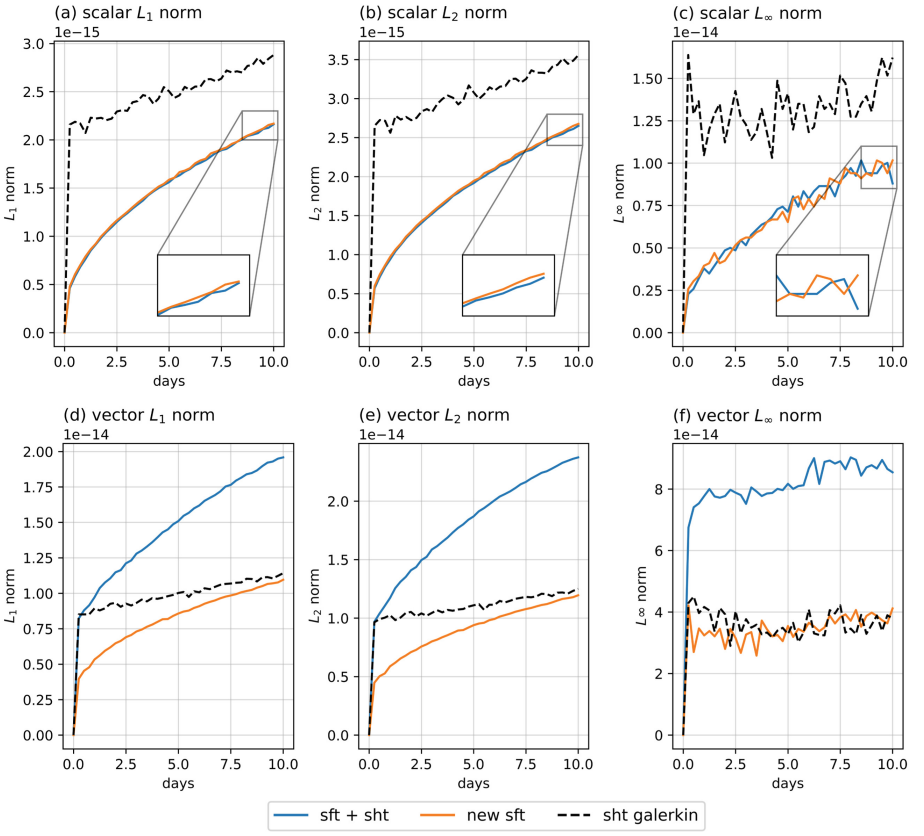


Fig. 6. Same as Fig. 3 but for Test Case 2. (a–c) scalar error norm, (d–g) vector error norm.

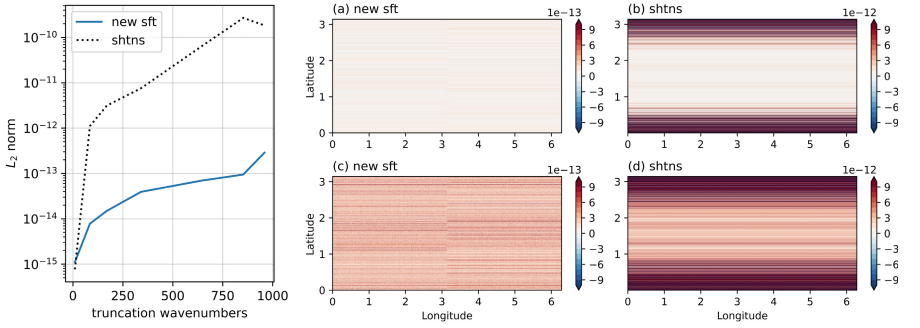


Fig. 7. Error of the proposed method and SHTns. Left panel shows the L_2 error norms. Right panel shows the error distribution in the physical space, (a,c) the proposed method and (b,d) SHTns. The truncation wavenumbers are (a,b) 639, (c,d) 959.

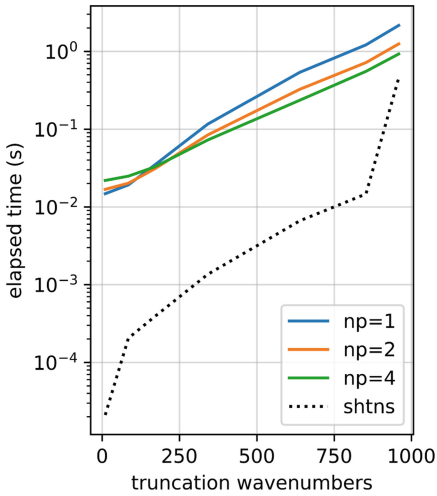


Fig. 8. Execution time for the proposed method and SHTns. Abscissa represents the numbers of truncation wavenumbers, and ordinates represents the execution time. np is parallelization number in Coarray Fortran. shtns parallelized with OpenMP using 64 threads.

The proposed method was parallelized using Fortran Coarrays, which were introduced in the Fortran 2008 standard. In contrast, the SHTns was compiled with OpenMP support and executed on 64 threads. The computational environment consisted of FreeBSD 14.3-RELEASE-p3 equipped with an AMD Ryzen Threadripper PRO 3975WX CPU (32 cores, 64 threads) and 256 GB of RAM.

For the benchmark test, we evaluated the computation of a gradient from a scalar field followed by wavenumber truncation. The input scalar was defined as follows:

$$f(\theta, \lambda) = \cos(2\theta) \tag{16}$$

The spherical Fourier transform method employed a regular latitude-longitude grid that includes the poles, whereas the spherical harmonic transform method utilized a Gaussian grid. Additionally, because SHTns performs the spherical harmonic transform at a predefined resolution, explicit wavenumber truncation is unnecessary.

First, we compare only the latitudinal component of the gradient. The error in the method using the spherical Fourier transform is approximately $O(10^{-2})$ smaller than that of the method using SHTns, indicating higher accuracy. In particular, a significant difference is observed in the error distribution near the poles (Fig. 7). SHTns employs a three-term recurrence relation to compute the associated Legendre functions [13,15]. At high resolutions, this method is prone to numerical underflow near the poles. Furthermore, it has been pointed out that computing the associated Legendre functions using a three-term recurrence relation yields lower accuracy compared to using a four-term recurrence relation

[6,7]. In contrast, the spherical Fourier transform maintains numerical stability and high accuracy even at high resolutions.

Next, we evaluate the computational performance by comparing the execution times of the two methods. To minimize the impact of background system processes, the computation was repeated 100 times, and the minimum execution time recorded across all trials is presented in Fig. 8.

The Coarray method was not faster than the SHTns method at resolutions with truncation wavenumbers of 1000 or less. However, as the excution time for the SHTns method increases sharply when the truncation wavenumber exceeds approximately 800, the spherical Fourier transform method may offer faster computation for high-resolution calculations. In the present study, the verification was limited to resolutions up to truncation wavenumbers of 1000 due to available computational resources and current implementation constraints. Performance evaluation at higher resolutions remain as a subject for future work.

6 Conclusions

We have developed a spherical Fourier transform method that achieves faster computation compared to the spherical harmonic transform. While numerical challenges remain in calculating vector quantities due to the complexity of the intermediate transformations, the proposed method achieves higher accuracy for scalar fields. This method possesses two novel features absent in previously proposed spherical Fourier transform approaches:

- The problem of the pole singularity is resolved by applying a specific formulation at the poler points, distinct from that used for non-poler grids. This enables the method to accommodate regular latitude-longitude grids that include the poles.
- By employing an isotropic wavenumber truncation on the sphere constructed solely with Fourier transforms, the method enables numerically stable and highly accurate computations.

These features facilitate the efficient solution of time-dependent partial differential equations on the sphere. Furthermore, because both of these features can be formulated without modifying the basis of the Fourier transform, there is no need to solve a system of linear equations to transform the coefficients. Consequently, the method can be implemented simply by utilizing any standard library capable of performing forward and inverse one-dimensional Fourier transforms.

Acknowledgments. The authors thank three anonymous reviewers for their constructive comments. This work was supported by JSPS KAKENHI Grant Number 24H00021 and JST K Program Grant Number JPMJKP25Q1, and was conducted as part of the collaborative research program (2026NE-01) of the Disaster Prevention Research Institute, Kyoto University.

References

1. Adams, J.C., Swarztrauber, P.N.: SPHEREPACK 3.0: a model development facility. Mon. Wea. Rev. **127**(8), 1872–1878 (1999). https://doi.org/10.1175/1520-0493(1999)127<1872:SAMDF>2.0.CO;2
2. Boyd, J.P.: The choice of spectral functions on a sphere for boundary and eigenvalue problems: a comparison of chebyshev, fourier and associated legendre expansions. Mon. Wea. Rev. **106**(8), 1184–1191 (1978). https://doi.org/10.1175/1520-0493(1978)106<1184:TCOSFO>2.0.CO;2
3. Cheong, H.B.: Application of double Fourier series to the shallow-water equations on a sphere. J. Comput. Phys. **165**(1), 261–287 (2000). https://doi.org/10.1006/jcph.2000.6615
4. Cheong, H.B.: Double Fourier series on a sphere: applications to elliptic and vorticity equations. J. Comput. Phys. **157**(1), 327–349 (2000). https://doi.org/10.1006/jcph.1999.6385
5. Cooley, J.W., Tukey, J.W.: An algorithm for the machine calculation of complex Fourier series. Math. Comp. **19**(90), 297–301 (1965). https://doi.org/10.2307/2003354
6. Enomoto, T.: Comparison of computational methods of associated Legendre functions. SOLA **11**, 144–149 (2015). https://doi.org/10.2151/sola.2015-033
7. Enomoto, T., Kuwano-Yoshida, A., Komori, N., Ohfuchi, W.: Description of AFES 2: improvements for high-resolution and coupled simulations. In: High Resolution Numerical Modelling of the Atmosphere and Ocean, pp. 77–97. Springer, New York, NY (2008). https://doi.org/10.1007/978-0-387-49791-4_5
8. Frigo, M., Johnson, S.: The design and implementation of FFTW3. Proc.IEEE **93**(2), 216–231 (2005). https://doi.org/10.1109/JPROC.2004.840301
9. Fukushima, T.: Fast computation of sine/cosine series coefficients of associated Legendre function of arbitrary high degree and order. J. Geod. Sci. **8**(1), 162–173 (2018). https://doi.org/10.1515/jogs-2018-0017
10. Fukushima, T.: Transformation between surface spherical harmonic expansion of arbitrary high degree and order and double Fourier series on sphere. J. Geod. **92**(2), 123–130 (2018). https://doi.org/10.1007/s00190-017-1049-3
11. Gruber, C., Abrykosov, O.: On computation and use of Fourier coefficients for associated legendre functions. J. Geodesy **90**(6), 525–535 (2016). https://doi.org/10.1007/s00190-016-0891-z
12. Hofsommer, D.J., Potters, M.L.: Table of fourier coefficients of associated legendre functions : repr. from Proceedings of the KNAW, ser. A. Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen: Series A: Mathematical Sciences **63**(5), 460–480 (1960). https://ir.cwi.nl/pub/8862
13. Ishioka, K.: A new recurrence formula for efficient computation of spherical harmonic transform. J. Meteor. Soc. Japan **96**(2), 241–249 (2018). https://doi.org/10.2151/jmsj.2018-019
14. Orszag, S.A.: Fourier series on spheres. Mon. Wea. Rev. **102**(1), 56–75 (1974). https://doi.org/10.1175/1520-0493(1974)102<0056:FSOS>2.0.CO;2
15. Schaeffer, N.: Efficient spherical harmonic transforms aimed at pseudospectral numerical simulations. Geochem. Geophys. Geosyst. **14**(3), 751–758 (2013). https://doi.org/10.1002/ggge.20071
16. Schuster, A.: On some definite integrals, and a new method of reducing a function of spherical co-ordinates to a series of spherical harmonics. Philos. Trans. A Math Phys. Eng. Sci. **200**(321–330), 181–223 (1903). https://doi.org/10.1098/rsta.1903.0004

17. Spotz, W.F., Taylor, M.A., Swarztrauber, P.N.: Fast shallow-water equation solvers in latitude-longitude coordinates. J. Comput. Phys. **145**(1), 432–444 (1998). https://doi.org/10.1006/jcph.1998.6026
18. Swarztrauber, P.N.: On the spectral approximation of discrete scalar and vector functions on the sphere. SIAM J. Numer. Anal. **16**(6), 934–949 (1979). https://doi.org/10.1137/0716069
19. Swarztrauber, P.N.: Vectorizing the FFTs. In: Parallel Computations, pp. 51–83. Academic Press (1982). https://doi.org/10.1016/B978-0-12-592101-5.50007-5
20. Swarztrauber, P.N.: The vector harmonic transform method for solving partial differential equations in spherical geometry **121**, 3415–3437 (1993). https://doi.org/10.1175/1520-0493(1993)121<3415:TVHTMF>2.0.CO;2
21. Williamson, D.L., Drake, J.B., Hack, J.J., Jakob, R., Swarztrauber, P.N.: A standard test set for numerical approximations to the shallow water equations in spherical geometry. J. Comput. Phys. **102**(1), 211–224 (1992). https://doi.org/10.1016/S0021-9991(05)80016-6
22. Yee, S.Y.K.: Studies on Fourier series on spheres. Mon. Wea. Rev. **108**(5), 676–678 (1980). https://doi.org/10.1175/1520-0493(1980)108<0676:SOFSOS>2.0.CO;2
23. Yee, S.Y.K.: Solution of Poisson's equation on a sphere by truncated double Fourier series. Mon. Wea. Rev. **109**(3), 501–505 (1981). https://doi.org/10.1175/1520-0493(1981)109<0501:SOPEOA>2.0.CO;2
24. Yoshimura, H.: Improved double Fourier series on a sphere and its application to a semi-implicit semi-Lagrangian shallow-water model. Geosci. Model Dev. **15**(6), 2561–2597 (2022). https://doi.org/10.5194/gmd-15-2561-2022

Explainable and Accurate Valvular Heart Disease Screening Using Spectrogram-Based 2D-CNNs on PCG Signals

Talha Ahmed Siddiqui[1], Sania Akhtar[1,2(✉)], Muhammad Taha Mansoor[1], Badee Baltow[3], Sadaq-Ur Rehman[4(✉)], and Noor Ul Huda[1]

[1] Faculty of Computer Science and Artificial Intelligence, Department of Creative Technology, Air University, Islamabad, Pakistan

[2] Faculty of Computer Science and Engineering, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology (GIKI), Topi 23460, Pakistan
sania.akhtar@au.edu.pk

[3] King Abdullah Medical City, Makkah, Saudi Arabia

[4] School of Sciences, Engineering and Environment, University of Salford, Manchester M5, UK
s.rehman15@salford.ac.uk

Abstract. Despite the recent advancements in artificial intelligence the cardiovascular diseases remain one of the leading causes of mortality worldwide, emphasizing the need for reliable and accessible diagnostic techniques. For the detection of the cardiac abnormalities, phonocardiography (PCG) which records heart sounds during the cardiac cycle, provides a non-invasive and cost-effective method. Conversely, manual interpretation of PCG signals depends heavily on clinician expertise and is often difficult by co-occurring pathologies, signal noise, and class imbalance in clinical datasets. This research focuses on an end-to-end deep learning pipeline for automated multi-label analysis of valvular heart diseases using PCG recordings. BMD-HS dataset containing 872 recordings sampled at 4000 Hz used for experimentations. The developed system implements a dual-stream learning strategy: handcrafted feature extraction combined with XGBoost for interpretability, and deep learning representations using one-dimensional and two-dimensional convolutional neural networks for improved predictive performance. Experimental results establish that altering raw PCG signals into Mel-Spectrogram representations and then it processes them with a 2D-CNN architecture achieving the best performance, with an AUC of 0.92 and an F1-score of 0.94 for the normal class. Furthermore, Explainable Artificial Intelligence techniques are implemented to confirm that the model focuses on medically significant spectral features. The developed framework provides a robust and low-cost solution for automated cardiac screening in resource-restricted healthcare environments.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 135–152, 2027.
https://doi.org/10.1007/978-3-032-30539-8_9

Keywords: Phonocardiography · Valvular heart disease detection · CNN · Mel-spectrogram analysis · Multi-label classification · Explainable artificial intelligence

1 Introduction

Cardiovascular diseases (CVDs) remain one of the leading causes of death globally, necessitating reliable and affordable diagnostic methods [1,2]. Among these, phonocardiography (PCG) captures acoustic signals produced by cardiac valve activity and associated hemodynamics [3,4]. In a healthy cardiac cycle, two primary sounds are observed: S1, generated by mitral and tricuspid valve closure at systole onset, and S2, generated by aortic and pulmonary valve closure at diastole onset. Valvular heart diseases manifest as murmurs arising from turbulent blood flow, exhibiting distinct variations in intensity, timing, and frequency [5,6]. Despite being non-invasive and cost-effective, classical auscultation is inherently subjective, relying heavily on clinician expertise, and is further compromised by environmental noise and physiological variability [7,8]. These limitations have driven growing interest in automated computational methods for PCG analysis.

Deep learning has significantly advanced automated biomedical signal analysis, with CNNs and hybrid architectures successfully applied to murmur detection, heart sound classification, and cardiac pathology identification from raw signals or time-frequency representations [9–13]. Alkhodari and Fraiwan [18] proposed a CNN-BiLSTM architecture achieving an AUC of 0.998 for PCG-based VHD detection, while Jamil and Roy [19] introduced a Vision Transformer (ViT) framework achieving 99.90% accuracy and 99.95% F1-score on time-frequency PCG representations. However, key challenges persist: PCG signals are non-stationary and noise-prone [14]; datasets frequently exhibit class imbalance [15]; and clinical adoption demands model interpretability [16]. Furthermore, most existing studies address binary or single-pathology classification and adopt a single representation strategy without systematic comparison across signal modalities [17].

To address these gaps, this study proposes a dual-stream diagnostic pipeline integrating interpretable handcrafted feature analysis with deep spectrogram-based representation learning, evaluated on the clinically realistic BMD-HS dataset. A clinically grounded explainability framework is incorporated to identify physiologically meaningful spectral regions associated with pathological murmurs, supporting future deployment in AI-assisted digital stethoscope systems for affordable cardiac screening in resource-constrained settings.

1.1 Contributions

The main contributions of this work are:

– A multi-label framework is proposed for automated screening of co-occurring valvular heart diseases using the publicly available BMD-HS dataset under clinically realistic noise and recording variability.

- A systematic comparative analysis is conducted across handcrafted acoustic descriptors, raw waveform-based 1D-CNNs, and Mel-spectrogram-based 2D-CNNs within a unified experimental framework.
- A spectrogram-driven 2D-CNN achieves an AUC of 0.92 and a Normal-class F1-score of 0.94, demonstrating strong discrimination between pathological and healthy heart sounds.
- A clinically interpretable explainability framework combining feature-importance analysis and Grad-CAM-based spectral localization confirms model predictions are driven by medically relevant murmur frequency regions.
- The framework supports future deployment in AI-assisted digital stethoscope systems for low-cost cardiac screening in resource-constrained environments.

2 Data Description and Exploratory Data Analysis (EDA)

2.1 Dataset Overview

The BMD-HS (Bangladesh Medical College – Heart Sounds) dataset used for this study, comprising of phonocardiogram (PCG) recordings collected from the National Institute of Cardiovascular Diseases, Dhaka. It comprised of 108 patients, each contributing up to 8 recordings from diverse auscultation points (Mitral, Tricuspid, Pulmonary, Aortic) and postures (Supine, Sitting), adding almost 872 audio files. All recordings are sampled at 4000 Hz, enough to capture vital high-frequency murmurs (up to 600–700 Hz) and heart sounds (20–200 Hz). Digital stethoscopes is used to capture recordings in a real-world clinical environment, which results in, speech artifacts, ambient noise, and sensor friction.

2.2 Multi-label and Class Imbalance Characteristics

A key feature of the dataset is its multi-label nature: patients may exhibit multiple co-occurring valvular pathologies e.g., mitral stenosis with aortic regurgitation. Figure 1 represent the class distribution of BMD-HS dataset.

- Pathological Cases (Sick): Roughly balanced counts among Aortic Stenosis (37), Aortic Regurgitation (43), Mitral Regurgitation (38), and Mitral Stenosis (38).
- Normal Cases (Healthy): Only 21 subjects labeled strictly as Normal (20% of the dataset).
- Implication: The scarcity of normal samples increases the risk of over-diagnosis. To address this, data augmentation techniques (pitch shifting, noise injection) are applied to the Normal class during training.

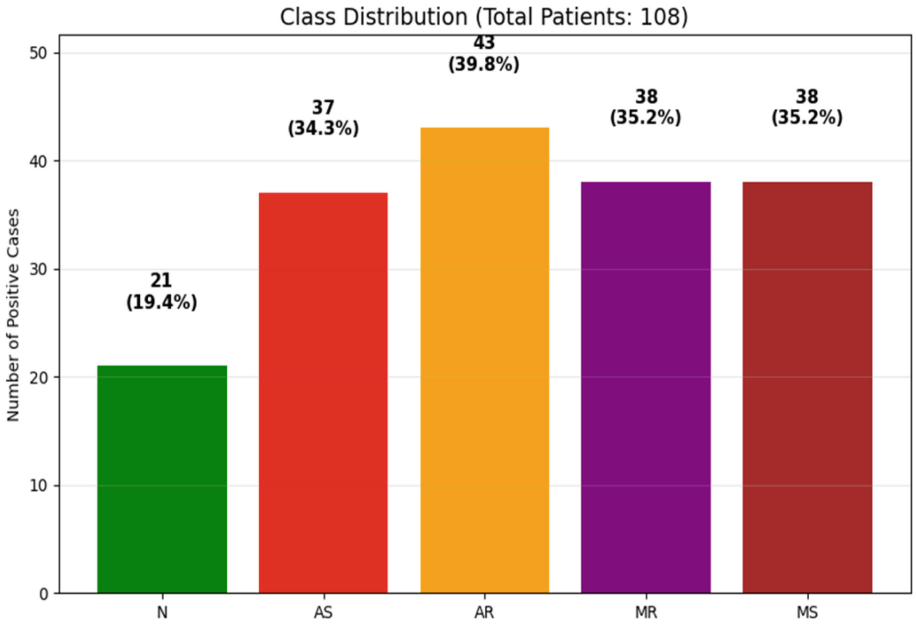


Fig. 1. The demonstration of class distribution of heart sound categories in the BMD-HS dataset comprising 108 patients.

2.3 Signal Characteristics

PCG signals are vary across patients and non-stationary. As represented in Fig. 2, the time-domain demonstration shows deviations in temporal structure and amplitude across cardiac cycles:

- Time: Evenly 20 s per recording.
- Amplitude: Active range differs due to stethoscope pressure and patient analysis. Signals are Z-score normalized to $[-1, 1]$.
- Noise Profile: Graphical examination exposes high-frequency sensor hiss and respiratory noise .

2.4 Spectral Analysis (Frequency Domain)

To examine spectral variances, time–frequency investigation was implemented using spectrogram representations and FFT, as presented in the right half of Fig. 2.

- Normal Signals: Energy determined in 20–200 Hz; S1 and S2 peaks are distinct and sharp, with clear quiet periods.
- Pathological Signals: Quiet intermissions interchanged by continuous noise (murmurs); substantial energy leakage in 200–600 Hz, emphasizing high-frequency modules as discriminative structures.

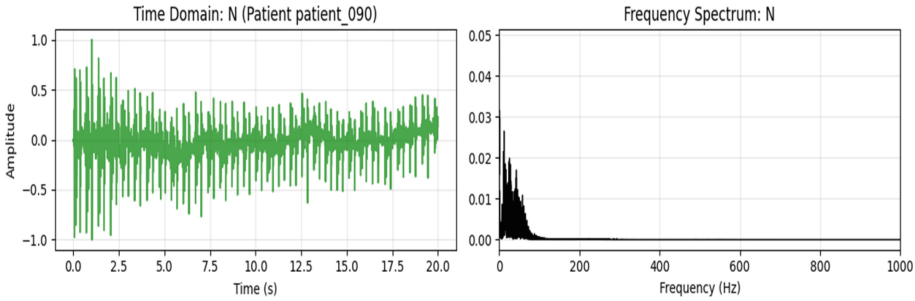


Fig. 2. Typical time-domain waveforms and their respective frequency-domain representations of PCG signals.

3 Methodology

The raw phonocardiogram (PCG) transforms recordings into automated diagnostic predictions through a proposed structured pipeline consisting of three main stages: signal preprocessing, data augmentation, and feature extraction and finally modeling. The suggested method enables the system to handle noisy real-world recordings while extracting significant acoustic features connected with cardiac irregularities. A advanced summary of the offered pipeline is demonstrated in Fig. 3.

3.1 Signal Preprocessing

The acquired PCG recordings in clinical environments frequently comprise of noise from ambient sounds, respiration, sensor friction and patient movement. A uniform preprocessing pipeline was applied to all recordings to enhance signal excellence and ensure robust feature extraction.

Resampling: At the native sampling rate of 4000 Hz all recordings were preprocessed. Sustaining this rate preserves significant high-frequency modules accompanying with pathological murmurs though escaping excessive loss of spectral evidence.

Bandpass Filtering: A second-order Butterworth bandpass filter with cutoff frequencies of 25 Hz and 400 Hz was applied. The ultimate heart sounds (S1 and S2) are typically focused below 150 Hz, while pathological murmurs may spread to higher frequencies. Frequencies below 25 Hz resemble mainly to standard wander produced by respiration, while frequencies above 400 Hz are mainly controlled by sensor noise.

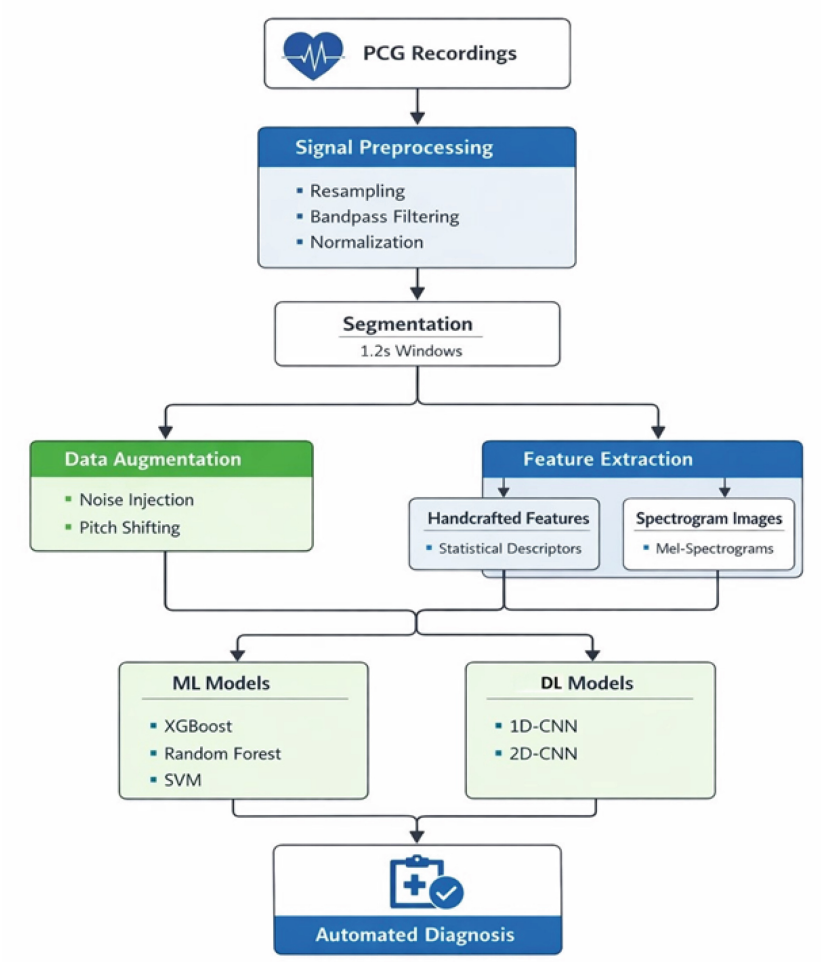


Fig. 3. The demonstration of workflow of proposed PCG-based automated diagnosis framework.

Amplitude Normalization: The recording volume variations caused by different stethoscope placements and patient physiology, signals using Min–Max scaling to normalize to the range [-1,1]. This guarantees that methods learn from waveform morphology rather than amplitude deviations.

Segmentation: The 20-second record was cut into 1.2-second pieces with an overlap of 50%. Medical scientists imagine a typical heart cycle at a rate of 60–100 beats per minute to be from 0.6 to 1.0 s. In this way, the window size chosen ensures that each segment of data has at least one full heartbeat.

3.2 Data Augmentation

There was a significant class imbalance in dataset, with only approximately 20% of normal subjects. To resolve this problem, data augmentation was applied specifically to the normal category during training. Two augmentation methods were used:

- **Noise Injection:** To simulate the low quality recordings Gaussian noise ($\mu = 0$, $\sigma = 0.005$) was added.
- **Pitch Shifting:** To emulate anatomical variations in heart size and acoustic propagation, signals were shifted by ± 2 semitones .

The applied data agumentation techniques increase the significant number of normal training samples. It also preserves the underlying acoustic characteristics of healthy heart sounds.

3.3 Feature Extraction

To leverage both the interpretability of classical machine learning and the pattern-recognition capabilities of deep learning, a dual feature extraction strategy was adopted. This approach extracts handcrafted statistical features for conventional machine learning models and deep visual features for convolutional neural networks.

Handcrafted Features: Handcrafted features capture mathematically defined properties of PCG signals and provide interpretable insights into acoustic characteristics associated with valvular abnormalities, as illustrated in Fig. 4.

Time-Domain Features.

- **Statistical Moments:** Mean, variance, skewness, and kurtosis describe the probability distribution of signal amplitudes.
- **Energy Metrics:** Root Mean Square (RMS) energy represents signal loudness.
- **Zero-Crossing Rate (ZCR):** Measures the frequency of sign changes in the signal and acts as a proxy for frequency content.
- **Crest Factor:** Ratio between peak amplitude and RMS value, indicating the impulsiveness of the waveform.

Frequency-Domain Features.

- **Mel-Frequency Cepstral Coefficients (MFCCs):** The first 13 coefficients were extracted to represent short-term spectral characteristics of the signal.
- **Spectral Centroid:** Indicates the center of mass of the spectrum and correlates with perceived brightness of the sound.
- **Spectral Entropy:** Measures the disorder of the spectral energy distribution.

Time–Frequency Features (Wavelet Analysis): Since PCG signals are non-stationary, wavelet-based analysis was used to capture localized time–frequency characteristics.

- A Discrete Wavelet Transform (DWT) using the Daubechies 4 (db4) mother wavelet was applied.
- The signal was decomposed into four levels.
- Energy and entropy were computed for both approximation and detail coefficients.

Feature Selection: To avoid the curse of dimensionality and remove redundant information, Principal Component Analysis (PCA) was applied to the aggregated feature set while retaining 99% of the cumulative variance.

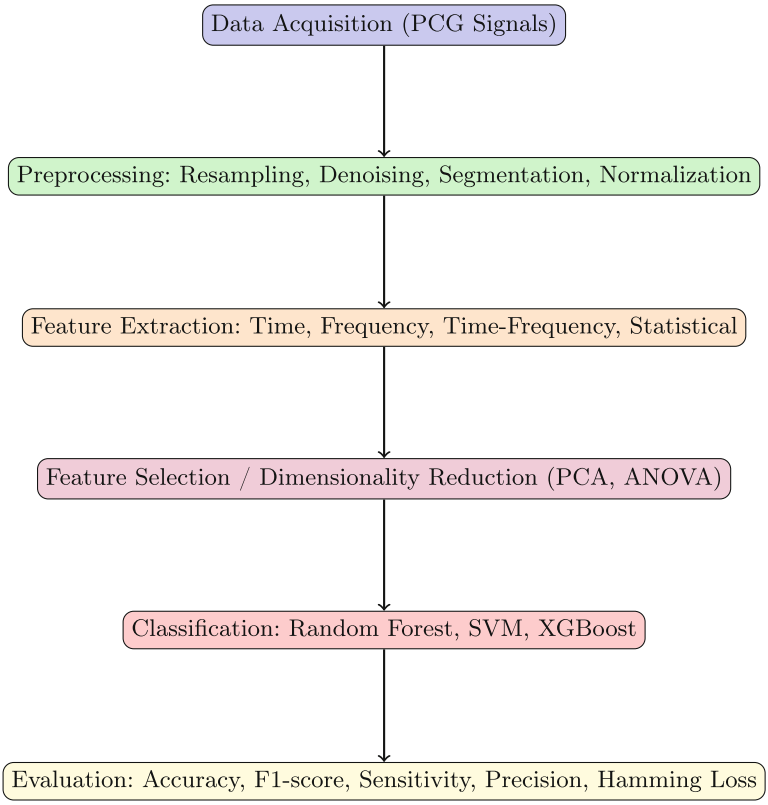


Fig. 4. A specific feature extraction pipeline carefully developed for classifying PCG signals.

Deep Visual Features. For convolutional neural network (CNN) models, the PCG classification problem was reformulated as a computer vision task by converting one-dimensional audio signals into two-dimensional time–frequency representations.

3.4 Mel-Spectrogram Generation Parameters

- Sampling rate: 4000 Hz
- FFT window size: 256 samples (∼64 ms)
- Hop length: 64 samples (∼16 ms overlap)
- Number of Mel bands: 64

The resulting power spectrogram was converted to the decibel (dB) scale to reflect the logarithmic sensitivity of human hearing. Each 1.2-second segment was therefore transformed into a 64 × 75 pixel Mel-spectrogram image, enabling CNN models to learn discriminative temporal–spectral patterns associated with cardiac abnormalities (Fig. 5).

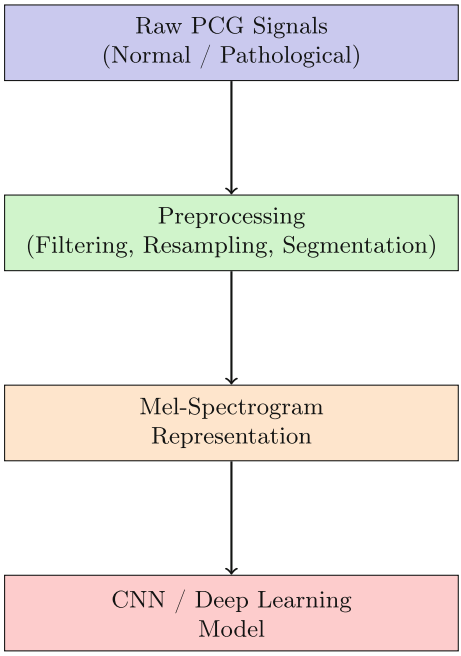


Fig. 5. The Pipeline represents the mel-Spectrogram for normal and pathological PCG Segments.

3.5 Modeling Strategy

Two complementary modeling strategies were employed.

- **Classical Machine Learning:** Handcrafted feature vectors were used to train interpretable models including XGBoost, Random Forest, and Support Vector Machines (SVM).
- **Deep Learning Models:** Spectrogram images were processed using 1D-CNN (raw waveform analysis) and 2D-CNN architectures operating on Mel-spectrogram representations.

This dual-stream framework combines the interpretability of statistical features with the representation learning capability of deep neural networks, enabling robust and clinically meaningful predictions.

4 Model Architectures and Comparative Analysis

To comprehensively evaluate different learning paradigms for heart sound classification, a diverse set of models was implemented. These models range from classical machine learning algorithms to deep neural networks and differ primarily in the type of input representation they process. Specifically, the models are grouped into three categories:

- **Feature-Based Learning (Tabular Data)** – models trained on handcrafted statistical, spectral, and wavelet features.
- **Time-Domain Representation Learning** – deep networks trained directly on raw PCG waveforms.
- **Time–Frequency Representation Learning** – convolutional models trained on Mel-spectrogram images.

This multi-model strategy enables a systematic comparison between interpretable feature-driven methods and deep representation learning approaches.

4.1 Feature-Based Learning (Tabular Data)

The first category of models utilizes the handcrafted feature vector derived from statistical, spectral, and wavelet features. After extraction, the feature space was reduced using Principal Component Analysis (PCA) to remove redundancy while retaining 99% of the variance. These models assume that the engineered features sufficiently capture the acoustic signatures associated with cardiac abnormalities.

4.2 Classical Machine Learning Classifiers:

Three widely used machine learning algorithms were evaluated as baseline models.

Random Forest (RF): Random Forest is an ensemble learning method consisting of multiple decision trees trained using the bagging strategy. In this study, a forest of **200 decision trees** was used. Random Forest was selected due to its robustness to noise and its ability to model complex non-linear interactions between features, such as relationships between wavelet energy and zero-crossing rate.

Support Vector Machine (SVM): Support Vector Machines were implemented with a **Radial Basis Function (RBF) kernel**. For the identification of optimal hyperplane that maximizes the margin between classes SVMs serves as high-dimensional feature spaces. This makes them perfect for discriminating subtle differences between pathological PCG and normal signals.

XGBoost (Extreme Gradient Boosting): To address the class imbalance in the dataset, the `scale_pos_weight` parameter was used to penalize misclassification of minority class samples. XGBoost was used for this purpose which is a gradient boosting algorithm that builds decision trees sequentially, where each new tree attempts to correct the errors of the previous ones.

Multi-layer Perceptron (MLP): A fully connected neural network was trained on the tabular feature vectors to bridge the gap between classical machine learning and deep learning.

- Input layer conforming to the PCA-reduced feature vector
- First hidden layer: 128 neurons with ReLU activation
- Second hidden layer: 64 neurons with ReLU activation
- Dropout layers (rate = 0.3)
- Output layer with sigmoid activation for classification

This architecture allows the model to pick up non-linear relationships between structures while preserving comparatively low computational complexity.

4.3 Time-Domain Representation Learning (1D-CNN)

A 1D Convolutional Neural Network (1D-CNN) was developed to learn discriminative features directly from the raw PCG waveform. Each input sample comprised of a raw audio segment with a length of 4800 samples, corresponding to $1.2\,s$ of audio recorded at a sampling rate of $4000\,Hz$. The proposed ID-CNN was developed to learn hierarchical temporal demonstrations from the PCG signals.

Initial convolutional layers operate with larger kernels (size = 9) to the low-frequency temporal patterns, for example, the rhythm and periodicity related to the S1 and S2 heart sounds. Whereas for the detection of high-frequency transient patterns deeper convolutional layers employ smaller kernels (size = 3). A GlobalAveragePooling1D layer was applied before the final classification stage instead of traditional fully connected layers to reduce the model complexity and improve generalization. This procedure breakdown the temporal dimension by computing the average activation of each feature map, by this means meaningfully dropping the number of trainable parameters and decreasing the risk of overfitting, which is particularly important when working with moderate-sized biomedical datasets.

4.4 Time–Frequency Representation Learning (2D-CNN)

The Mel-spectrogram images used as a input for convolutional neural network for classification and this is the third modeling approach. In this approach, each PCG segment is converted into a 64 × 75 Mel-spectrogram image , which demonstrats the time–frequency energy distribution of the heart sound signal. A VGG-19 CNN based architecture was implemented to process these spectrogram representations. The model comprised of three blocks, Conv2D layer, followed by Batch Normalization, a ReLU activation function, and a MaxPooling2D layer. These arranged convolutional blocks gradually extract higher-level spectral features from the Mel-spectrogram images. The 2D convolution filters operate across both the time axis (horizontal dimension) and the frequency axis (vertical dimension) unlike 1D-CNN models that analyze signals purely in the temporal domain. This allows the network to learn spatiotemporal features within the spectrograms, permitting it to recognize visually diverse acoustic features such as broadband noise areas related with murmurs and transient spectral peaks equivalent to valve clicks.

4.5 Training Configuration

A consistent training configuration was adopted to ensure a fair comparison across all neural network models (MLP, 1D-CNN, and 2D-CNN). A Binary Cross-Entropy loss function and the Adam optimizer with a learning rate of 0.001 used for all model trainings. A ReduceLROnPlateau scheduler was employed to reduce the learning rate when the validation loss stopped improving, whereas early stopping was employed to dismiss training if the validation loss did not improve for 10 successive epochs. This standardized training protocol guaranteed that performance differences primarily returned architectural proficiencies relatively than optimization setting variations (Table 1).

5 Results and Discussion

The quantitative evaluation of the proposed models for heart sound classification is presentsed in this section. The 80/20 train-test split was performed at the

Table 1. Training Parameters

Parameter	Value/Description
Framework	Python, TensorFlow/Keras
Maximum Epochs	100
Batch Size	32
Early Stopping Patience	10 epochs (monitored on validation loss)
Train–Test Split	80/20 stratified split
Validation Set	Subset of training partition
Intermediate Activation	ReLU
Output Activation	Sigmoid
Loss Function	Binary Crossentropy
Optimizer	Adam
Learning Rate	0.001
LR Scheduling	ReduceLROnPlateau
Classification Type	Multi-label

patient level, ensuring that all segments derived from a given patient appeared exclusively in either the training or test set. This strict partitioning eliminates the risk of data leakage and ensures that the model is evaluated on entirely unseen patients rather than unseen segments from familiar subjects. The Model was evaluated on the held-out test set using metrics appropriate for imbalanced multi-label classification. In addition to classical machine learning baselines, the evaluation framework was designed to compare handcrafted feature learning, raw waveform representation learning, and spectrogram-based deep learning strategies in order to comprehensively assess the effectiveness of different PCG signal representations for valvular heart disease screening.

5.1 Evaluation Metrics

Traditional accuracy alone is insufficient for evaluating multi-label medical classification systems. A model predicting only the majority class may achieve high accuracy while failing to detect diseases. Therefore, multiple complementary metrics were used.

Exact Match Accuracy: measures the percentage of samples for which the predicted label set exactly matches the ground truth:

$$Accuracy = \frac{1}{N}\sum_{i=1}^{N} I(y_i = \hat{y}_i)$$

Since this metric is strict for multi-label problems, we primarily rely on the

Macro F1-Score: averages the F1-score across all classes:

$$Macro\ F1 = \frac{1}{L}\sum_{l=1}^{L}\frac{2P_l R_l}{P_l + R_l}$$

In addition, two clinically relevant metrics were considered:

Recall (Sensitivity): measures the ability to detect pathological cases.

Precision: measures the reliability of positive predictions.

Hamming Loss: was used to measure the fraction of incorrectly predicted labels:

$$HammingLoss = \frac{1}{N \times L}\sum_{i=1}^{N}\sum_{l=1}^{L} xor(y_{i,l}, \hat{y}_{i,l})$$

Lower Hamming Loss indicates better multi-label prediction performance.

5.2 Performance of Feature-Based Models

Table 2 summarizes the performance of models trained on the handcrafted feature vector consisting of statistical, spectral, and wavelet features. Classical machine learning models achieved relatively low exact-match accuracy, indicating difficulty in capturing complex dependencies between multiple cardiac conditions. Among them, XGBoost produced the strongest baseline performance with a Macro F1-score of 0.61, demonstrating its ability to detect pathological patterns in the feature space.

Table 2. Performance of feature-based models

Model	Accuracy	Macro F1	Hamming Loss
Random Forest	24.8%	0.49	0.22
SVM (RBF)	28.5%	0.52	0.20
XGBoost	35.2%	0.61	0.18
MLP	89.0%*	0.68	0.12

The Multi-Layer Perceptron (MLP) significantly improved performance, achieving a Macro F1-score of 0.68. This suggests that relationships among spectral and wavelet features are highly nonlinear and benefit from neural network representations.

5.3 Performance of Representation Learning Models

Deep learning models trained directly on signal representations demonstrated further improvements in classification performance. Table 3 compares the results of the 1D-CNN (time-domain) and 2D-CNN (spectrogram-based) architectures.

Table 3. Deep learning model comparison

Metric	1D-CNN	2D-CNN
AUC	0.80	0.92
Macro F1	0.59	0.76
Precision	0.58	0.76
Recall	0.77	0.81
Normal Class F1	0.73	0.94

The 1D-CNN demonstrated high recall for pathological classes, indicating strong sensitivity for detecting abnormal heart sounds. This marks it appropriate as a screening model. Though, the model demonstrated comparatively low precision due to difficulty differentiating noisy common signals from murmurs. Comparatively, the 2D-CNN trained on Mel-spectrogram illustrations attained the best overall presentation with an AUC of 0.92. The model considerably improved classification of normal cases, achieving a normal class F1-score of 0.94. These results propose that spectrogram demonstrations provide richer time–frequency information, which enables the CNN to learn acoustic features connected with cardiac murmurs more efficiently. Experimental observations showed that smaller FFT windows weakened the spectral discrimination of murmurs, while larger windows introduced temporal smoothing that hurt the localization of transient acoustic patterns.

5.4 Comparison with State-of-the-Art

Table 4. Comparison with existing state-of-the-art methods for valvular heart disease detection using PCG signals

Study	Method	Dataset	AUC	F1/Accuracy
Alkhodari & Fraiwan (2021)	CNN-BiLSTM	Private clinical	0.998	99.32% Acc
Jamil & Roy (2023)	Vision Transformer (ViT)	Private PCG	__	99.95 F1
Proposed	**2D-CNN (Mel-spectrogram)**	**BMD-HS**	**0.92**	**0.76 F1**

To contextualize the performance of the proposed framework, Table 4 presents a comparison with closely related state-of-the-art studies on PCG-based valvular heart disease detection. Alkhodari and Fraiwan [18] proposed

a CNN-BiLSTM architecture achieving an AUC of 0.998 on a private clinical dataset. Jamil and Roy [19] employed a Vision Transformer (ViT) with an accuracy of 99.90% and F1-score of 99.95% on a private PCG dataset. While these results appear higher, it is important to note that both studies operate on private datasets under binary or limited-class settings, whereas the proposed framework addresses the more challenging and clinically realistic problem of multi-label classification on the publicly available BMD-HS dataset.

5.5 Explainability Insights

Explainability techniques were applied to both feature-based and deep learning models to interpret model predictions. XGBoost model identified spectral centroid and wavelet energy as the most discriminative patterns. These conclusions align with clinical knowledge, as pathological murmurs naturally display amplified high-frequency energy due to turbulent blood flow. Additionally, Grad-CAM visualizations of the 2D-CNN discovered that the model emphases mainly on spectral regions between 200–600 Hz during systolic and diastolic intervals. This authorizes that the network learns medically appropriate acoustic features rather than background noise. It improves confidence in the model's diagnostic predictions. While Grad-CAM visualizations offered meaningful insight into the spectral regions driving model predictions, quantitative clinical validation of explainability remains an important area for future work. In this study, the highlighted regions were interpreted against established clinical characteristics of valvular murmurs from the cardiology literature, particularly the concentration of abnormal acoustic energy in higher-frequency bands. Future work will involve expert cardiologist evaluation and quantitative localization analysis to further validate the clinical reliability of the explainability framework.

6 Conclusion and Future Work

This research work proposed an end-to-end deep learning pipeline for the automated multi-label classification of valvular heart diseases using PCG recordings. The proposed system addresses key challenges in clinical datasets, including signal noise, co-occurring pathologies, and class imbalance by integrating handcrafted feature analysis with deep representation learning. Experimental results proves that the 2D-CNN architecture trained on Mel-Spectrogram demonstration gains the highest predictive performance, with an AUC of 0.92 and a Normal class F1-score of 0.94. The model effectively mitigates the over-diagnosis issues observed in 1D and classical feature-based models by treating heart sounds as visual textures.

– Clinical Significance: The model shows high clinical significance by maintaining high sensitivity for critical conditions such as Aortic Regurgitation (Recall $> 80\%$). It proves that its potential use as a low-cost, non-invasive screening tool in resource-constrained healthcare settings.

– Trustworthiness: Th model focuses on medically relevant spectral features with the help of post-hoc explainable AI analyses, ensuring that predictions are guided by physiologically meaningful patterns rather than dataset artifacts.

6.1 Future Work

Future research will focus on improving the accessibility, utility, and generalization of the proposed framework. One encouraging direction is to incorporate the 1D-CNN, which displays high sensitivity, with the 2D-CNN, which offers high specificity, into a weighted ensemble model. This will potentially optimize the balance between recall and precision for clinical deployment. Furthermore, model optimization and quantization for embedded platforms, such as Raspberry Pi devices or smartphone, could enable a standalone AI-powered stethoscope for point-of-care cardiac screening. Finally, framework evaluation on independent datasets, such as PhysioNet, will be essential to assess its generalizability across different recording devices and diverse patient populations. It ensures robust real-world performance. Overall, the proposed method provides a interpretable, robust, and low-cost solution for automated cardiac screening, which represents a significant step toward accessible AI-driven healthcare diagnostics.

References

1. Ayappasamy, K., et al.: Detection of valvular heart diseases from PCG signals using machine and deep learning models–a review. IEEE Access (2025)
2. Pavithra, G.R., et al.: Enhanced cardiovascular disease detection from phonocardiogram signals using deep learning & wavelet-based denoising. In: 2024 4th Asian Conference on Innovation in Technology (ASIANCON). IEEE (2024)
3. Muhammad Rafli, R., et al.: Enhanced identification of valvular heart diseases through selective phonocardiogram features driven by convolutional neural networks (SFD-CNN). Jurnal Nasional Teknik Elektro, 20–35 (2024)
4. Anjan, G., et al.: Automated system for the detection of heart anomalies using Phonocardiograms: a systematic review. IEEE Access **12**, 138399–138428 (2024)
5. Hend Khalid, A., et al.: Precision diagnosis: an automated method for detecting congenital heart diseases in children from phonocardiogram signals employing deep neural network. IEEE Access **12**, 76053–76064 (2024)
6. Zihan, J., et al.: Automated valvular heart disease detection using heart sound with a deep learning algorithm. IJC Heart Vasculature **51**, 101368 (2024)
7. Bollapalli, A., Panini Challa, N.: Explainable attention-based deep learning for classification and interpretation of heart murmurs using phonocardiograms. Sci. Reports **15**(1), 37991 (2025)
8. Subham Kumar, P., Mohapatra, A., Patra, S.: X-CBNet: an explainable effective deep learning framework based on spectrograms for predicting valvular disorder using PCG signals. J. Transform. Technol. Sustain. Dev. **9**(1), 18 (2025)
9. Popal Khan, P., et al.: Enhancing cardiac health diagnoses through machine learning analysis of phonocardiograms (PCG). J. Data Sci. Intell. Syst. **3**(4), 312–323 (2025)

10. Abdul, J., et al.: Automated detection of pediatric congenital heart disease from phonocardiograms using deep and handcrafted feature fusion. Comput. Biol. Med. **197**, 110993 (2025)
11. Hussain, S.S., Ashfaq, M., Khan, M.S., Anwar, S.: Deep learning based phonocardiogram signals analysis for cardiovascular abnormalities detection. In: 2023 International Conference on Robotics and Automation in Industry (ICRAI), Peshawar, Pakistan, pp. 1-6 (2023). https://doi.org/10.1109/ICRAI57502.2023.10089537
12. Ihtisham Ul, H., et al.: Enhancing pediatric congenital heart disease detection using customized 1D CNN algorithm and phonocardiogram signals. Heliyon, **11**(3) (2025)
13. Hemalatha, G., et al.: Precision diagnosis: an automated method for detecting congenital heart disease in children from phonocardiogram signals employing deep neural network. In: Recent Trends in Data Analytics and Computing, pp. 223–228. CRC Press (2024)
14. Orozco-Reyes, L., Alonso-Arévalo, M.A., García-Canseco, E., Ibarra-Hernández, R.F., Conte-Galván, R.: A deep-learning approach to heart sound classification based on combined time-frequency representations. Technologies **13**, 147 (2025). https://doi.org/10.3390/technologies13040147
15. Sukant, B., Saha Misra, I., Naveed Siddiqui, K.: Phonocardiogram signal based better prediction of aortic and mitral stenosis heart diseases using machine learning. In: 2025 International Conference on Ambient Intelligence in Health Care (ICAIHC). IEEE (2025)
16. Sukant, B., Saha Misra, I., Naveed Siddiqui, K.: Detection of aortic and mitral stenosis valvular heart diseases using machine learning on PCG signals. In: 2024 IEEE Calcutta Conference (CALCON). IEEE (2024)
17. Soh-Yeon, L., et al.: Deep learning-based evaluation of the severity of mitral regurgitation in canine myxomatous mitral valve disease patients using digital stethoscope recordings. BMC Veterinary Res. **21**(1), 326 (2025)
18. Alkhodari, M., Fraiwan, L.: Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings. Soft. Comput. **25**, 7017–7030 (2021)
19. Jamil, S., Roy, A.M.: An efficient and robust Phonocardiography (PCG)-based Valvular Heart Diseases (VHD) detection framework using Vision Transformer (ViT). Biomedical Signal Processing and Control (2023)

Revisiting Ray-Tracing Algorithm for Future Limb-Scanning Missions

Francesco Pio De Cosmo[1,2(✉)] and Luca Sgheri[2]

[1] Department of Information Engineering, University of Florence, Florence, Italy
[2] National Research Council, Institute for Applied Calculus, Sesto Fiorentino, Italy
francescopio.decosmo@unifi.it, luca.sgheri@fi.iac.cnr.it

Abstract. In the forward model for limb-scanning instruments, ray tracing must be accounted for because variations in air refractivity cause the lines of sight to bend from straight paths into curves. The tangent point of a line of sight defined as the minimum height, depends on both the instrument's nadir angle and the atmospheric state. To achieve reliable tangent point determination, the off-nadir angles must be calibrated to account for Earth's ellipsoidal geometry and realistic atmospheric conditions sampled by these lines of sight paths. In this study, we improve the ray-tracing algorithm originally developed for MIPAS operational retrievals. The algorithm is applied to more accurately assess the impact of atmospheric variability on the relationship between off-nadir angles and tangent point localization. It also allows us to estimate the smoothing error associated with the coarse horizontal sampling of the MIPAS observation pattern. We used a configuration representative of the proposed CAIRT mission instrument as a case scenario for future instruments such as the STRIVE mission by NASA.

Keywords: Remote sensing · CAIRT · Ray tracing algorithm · Limb scanning

1 Introduction

Limb scanning is an observation geometry in which the instrument looks tangentially through the atmosphere, without intersecting the Earth's surface. Compared to nadir-viewing satellites that observe vertically toward the Earth's surface, limb sounders achieve much higher vertical resolution, offering more detailed information on the upper atmosphere. In contrast, nadir looking instrument, have limited sensitivity to changes in the upper layers composition, as the signal from higher altitudes is largely masked by the stronger contributions coming from the lower layers.

However, limb instruments face two problems. First, the mechanics of the instrument is normally more complicated with respect to a nadir looking instrument, because it needs more complex optics [4]. Second, from the point of view of the forward modeling and retrieval procedures, there is the additional complications due to predict the shape of the lines of sight, since the changes in the

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 153–170, 2027.
https://doi.org/10.1007/978-3-032-30539-8_10

atmosphere refractive index bend the rays from straight lines into curves. This makes it necessary to accurately determine the tangent points, where the line of sight is closest to the Earth, in order to correctly assign altitudes to the retrieved atmospheric profiles.

Only a limited number of limb-scanning instruments have operated in space. On ESA's ENVISAT mission (2002–2012), SCIAMACHY (*Scanning Imaging Absorption Spectrometer for Atmospheric Cartography*) [5] and MIPAS (*Michelson Interferometer for Passive Atmospheric Sounding*) [12] provided valuable measurements, with MIPAS delivering geolocated profiles of temperature, pressure, and up to 20 atmospheric gases from 6 to 60 km altitude [8,21]. Other notable instruments include ACE-FTS (*Atmospheric Chemistry Experiment – Fourier Transform Spectrometer*) on Canada's SCISAT (2003) [2] and MLS (*Microwave Limb Sounder*) on NASA's Aura (2004) [28], which have supplied high-quality vertical profiles of atmospheric composition, temperature, and dynamics in the last two decades. However, both ACE-FTS and MLS are now nearing the end of their operational lifetimes.

The projected CAIRT (*Changing-Atmosphere Infra-Red Tomography Explorer*) instrument [26] instrument was conceived as a limb-scanning instrument and a potential successor to MIPAS. Although it was not selected for the 11th Earth Explorer mission in the final User Consultation Meeting, CAIRT provides a valuable framework for testing methodological advances, both in hardware and in software, for future limb-scanning missions. CAIRT was designed to acquire continuous measurements along the satellite orbit, with each acquisition capturing a 2D set of limb-emission spectra via individual detector pixels. The vertical coverage extends from about 5 km up to 115 km, with a vertical resolution of about 1 km up to the mid-mesosphere and 2 km above. The horizontal coverage extends over a swath of about 500 km, with along-track sampling of about 50 km and across-track sampling of 50 km. Further details on the instrument characteristics are described in the Mission Report [10].

Recently, the STRIVE (Stratosphere Troposphere Response using Infrared Vertically-resolved light Explorer) mission [27] has been selected for continued development by NASA. The mission contains ALICE (Advanced Limb Infrared Chemistry Experiment), an infrared limb-scanner that has many similarities with the CAIRT project.

In the forward model for limb-scanning instruments, the ray tracing must be taken into account. The retrieved lines of sight, and thus the accuracy of the tangent point positions, depend on several factors: the numerical ray-tracing method, the atmospheric refractive index model, and the assumed atmospheric composition. As suggested by the work [24] for MIPAS, variations in the ray-tracing method or refractive index model produce negligible errors ($\lesssim$ 2 m), suggesting that the choice of atmospheric model may have a more significant effect on the accuracy of tangent point determination and, consequently, of the retrieved profiles.

In this paper, we extend this study by implementing an improved ray-tracing algorithm originally developed for MIPAS operational retrievals. Using this algo-

rithm we provide a more accurate analysis of the relationship between instrument off-nadir angles and tangent point altitudes. In particular, we focus on the effect of using real atmospheric refractivity data, including seasonal variations, rather than standard models, which may introduce significant uncertainties in tangent point determination. This approach also allows us to estimate the smoothing error associated with the coarse horizontal sampling of the MIPAS observation pattern, which can now be assessed using a finer atmospheric discretization.

In Sect. 2, we provide a brief overview of ray tracing and the key aspects involved. Section 3 presents a methodology for predicting the instrument off-nadir angles without a priori knowledge of the sampled atmosphere. In Sect. 4, we report the main results of our analysis about displacements in tangent point localization Subsect. 4.1. In Subsect. 4.2, we estimate the MIPAS smoothing error due to the instrument's coarse horizontal resolution as an additional application of the algorithm. Finally, in Sect. 5 we derive some conclusions and ideas for the path ahead. In Appendix A we provide a brief description of the satellite orbit and atmospheric grid, as well as the geometry of the lines of sight adopted in the simulations.

2 Ray Tracing

2.1 Atmospheric Refractivity Models

The line of sight (LoS) is the trajectory of an electromagnetic ray, governed by the atmospheric refractive index n. Since n exceeds unity only slightly, it is convenient to define the refractivity $N = n - 1$ to quantify this deviation. Ray propagation in such a medium is described by the Eikonal equation [3], which is derived from the wave equation under the approximation that n varies slowly with respect to the radiation wavelength [1], as is the case of the infrared band. The LoSs are then the solution of this equation, that can be written as:

$$\frac{d}{ds}\left(n\big(\mathbf{L}(s)\big)\frac{d\mathbf{L}(s)}{ds}\right) = \nabla n\big(\mathbf{L}(s)\big), \tag{1}$$

where s is the arc parameter and $\mathbf{L}(s)$ represents the LoS. The refractive index n is mainly governed by air density and can be computed using standard models of atmospheric refraction. The Edlén model [9] calculates the refractivity as

$$n - 1 = \left(\frac{\rho_a}{\rho_{a,\mathrm{st}}}\right) N_{a,\mathrm{st}}(\lambda), \tag{2}$$

where the refractivity depends linearly on the atmospheric dry air density ρ_a, scaled by a standard density $\rho_{a,\mathrm{st}}$ at sea-level conditions ($p_{\mathrm{st}} = 1013.25$ hPa, $T_{\mathrm{st}} = 273.15$ K). Here, $N_{a,\mathrm{st}}$ represents the specific refractivity of dry air at standard conditions and depends on the wavelength λ . However, this simple model does not account for the effect of water vapor.

Ciddor [6] provides a more accurate formulation that includes the dependence on water vapor content and CO_2 concentration. According to Mathar [17], the

Ciddor method achieves an accuracy of about 10^{-6} in the refractive index for wavelengths up to 25 mm, covering the full relevant infrared spectral range. The refractivity, calculated following the Ciddor formulation, is expressed as

$$n - 1 = \left(\frac{\rho_a}{\rho_{a,\mathrm{st}}}\right) N_{a,\mathrm{st}}(\lambda) + \left(\frac{\rho_w}{\rho_{w,\mathrm{st}}}\right) N_{w,\mathrm{st}}(\lambda), \tag{3}$$

where ρ_a and ρ_w are the actual densities of dry air and water vapor under the given atmospheric conditions, $\rho_{a,\mathrm{st}}$ and $\rho_{w,\mathrm{st}}$ correspond to these at standard temperature and pressure values and $N_{a,\mathrm{st}}$, $N_{w,\mathrm{st}}$ are the specific refractivities of dry air and water vapor at standard reference conditions and wavelength. In this formulation, $N_{a,\mathrm{st}}$ also introduces a dependence on the CO_2 concentration. For both methods we chose a representative mid-infrared wavelength 10 mm, and a CO_2 mixing ratio of 425 ppmv for the Ciddor method, that is the one that we adopt in this study.

2.2 Geometric Configuration and Numerical Approach

A common and reasonable assumption is that the satellite's orbit around Earth can be approximated as circular. This approximation is justified by the requirement for the satellite to maintain an almost constant altitude above the Earth's surface [11], so the orbit eccentricity must be low. Also, the orbit is chosen to be sun-synchronous, that ensures that the satellite passes over any given point on Earth at nearly the same local solar time. For the Earth's shape, we chose to adopt the WGS84 rotational ellipsoid [19]. A more detailed discussion of the geometric considerations is given in Appendix A. For the ray-tracing procedure, a model of the atmosphere must be selected, including at least pressure, temperature and water vapor as a function of the orbital coordinate and the altitude above the WGS84 surface. In the present study, two distinct models were considered:

- The *US Standard Atmosphere 1976* (US76) [20], representing a reference standard atmosphere that provides globally averaged conditions. This model atmosphere does not depend on the geolocation or time.
- The atmosphere from the *ERA5 reanalysis dataset* [13], representing a realistic atmosphere including profiles of pressure, temperature, water vapour, and CO_2 from CAMS [7], extended up to 120 km using the IG2 climatology [23]. This model is used to represent the *true* atmospheric state.

For the numerical implementation of the Eikonal Eq. (1), we adopt the Adaptive Eikonal (AEIK) method described in [24], which employs a multi-step predictor-corrector approach and an adaptive step size based on the curvature of the line of sight. This allows for larger step sizes (up to 1 km) in regions with low atmospheric density, while reducing the step size to as small as 10 m in denser layers. The LoS tracing is performed using a backward ray-tracing algorithm. Starting from the satellite position, propagated through the atmosphere according to the solution of the Eikonal equation, reaching the tangent point at

its minimum altitude, and subsequently followed until it exits the atmospheric domain. We revisited the MIPAS ray-tracing algorithm and we improved the numerical stability of the solution method.

3 Methods

3.1 Determination of the Off-Nadir Angles

The off-nadir angles, or pointing angles, are defined as the angles between the direction of the satellite detectors and the line from the satellite normal to the Earth's surface. These angles are determined a-priori to ensure that the lines of sight intersect a predefined set of target *engineering* tangent altitudes, which defines the instrument's viewing geometry. The relationship between off-nadir angles and tangent height, however, are influenced by the Earth eccentricity and the atmospheric composition sampled by the line of sight. The Earth curvature, at least the value predicted by the WGS84 model, is known and can therefore be accounted for in advance. On the other hand the actual atmospheric composition is only known after the retrieval operation. But the off-nadir angles must be predeterminated, requiring an approximation of the atmosphere, which naturally introduces a bias between the actual tangent altitudes reached and the target *engineering* tangent altitudes. This choice influences the magnitude of the bias, as it will be shown in Sect. 4.

To account for the Earth's ellipticity, the adjustment of the off-nadir angles depends on the instrument hardware. The MIPAS instrument used a sequential acquisition of the sweeps, so each off-nadir angle could be separately adjusted [8]. The CAIRT instrument foresaw a global pitch adjustment [10] to counter variations of tangent altitudes along the orbit. Other instrumental models might provide different adjustment techniques. In this study we assume to be able to determine each off-nadir angle separately. Based on this assumption, we now describe the method we adopted to account for the Earth shape, which can be treated a priori, before performing ray tracing. We only consider a single set of off-nadir angles, restricting the ray-tracing to the orbital plane, as already mentioned. This approach is consistent with the planned retrieval procedure for CAIRT, which operates independently for each along-track line of sight (Fig. 1).

The goal is to determine the off-nadir angles $\{\hat{\alpha}_i\}_{i=1,\ldots,N}$, where N corresponds to the number of detectors, that produce tangent altitudes as close as possible to $\{\hat{h}_i\}_{i=1,\ldots,N}$, taking into account variations in the Earth's ellipticity as a function of the satellite position during each acquisitions.

The strategy employed to determine $\{\hat{\alpha}_i\}$ from the sampled $\{\hat{h}_i\}$ is the following:

(i) We define a preliminary set of m candidate off-nadir angles $\{\alpha_i\}_{i=1,\ldots,m}$ sampled from the interval $[\alpha_{\min}, \alpha_{\max}]$, with $\alpha_{\min} = 61°$ and $\alpha_{\max} = 65°$, chosen conservatively to ensure that rays remain within the atmosphere and do not intersect the Earth's surface [10].

$$\alpha_1 = \alpha_{\min}, \ \ldots, \ \alpha_M = \alpha_{\max}. \tag{4}$$

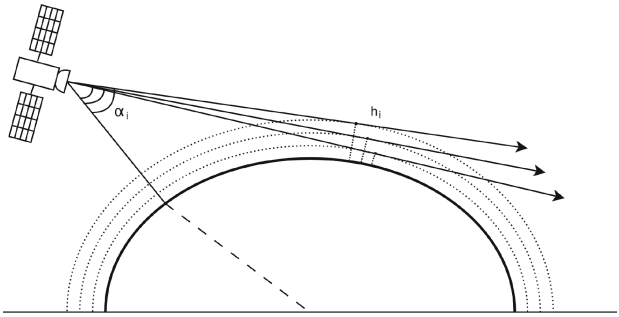


Fig. 1. Schematic of the satellite viewing geometry, showing off-nadir angles $\{\alpha_i\}$ and their corresponding tangent points $\{h_i\}$ for a fixed satellite acquisition.

This sampling does not need to be uniform and is independent of the particular satellite acquisition.

(ii) We define a set of satellite positions along the orbit, corresponding to sequential satellite acquisition and indexed by $j = 1, \ldots, N_{\text{acq}}$.

(iii) For each acquisition j, we trace the lines of sight $\mathbf{L}(s)_{ij}$ corresponding to the set $\{\alpha_i\}$, obtaining a new set of tangent altitudes candidate $\{h_{ij}\}$. We then define a linear spline $s_j(h)$ built using the nodes:

$$(h_{ij}, \alpha_i), \quad i = 1, \ldots, m \qquad j = 1, \ldots, N_{\text{acq}}. \tag{5}$$

(iv) Finally, for each acquisition j, evaluate the spline at the selected engineering tangent altitudes $\{\hat{h}_i\}$ to obtain the required off-nadir angles:

$$\hat{\alpha}_{ij} = s_j(\hat{h}_i), \quad i = 1, \ldots, N \qquad 1, \ldots, m \qquad j = 1, \ldots, N_{\text{acq}}. \tag{6}$$

3.2 Characterization of a Atmospheric Model

An atmospheric model is then required to apply the procedure described in the previous section. Two possible choices have been made.

- **Geometric model:** The lines of sight are computed using a straight-line approach, neglecting the bending effect of atmospheric refraction.
- **Standard US76 model:** The lines of sight are computed assuming a standard US76 atmosphere, a globally uniform atmospheric model.

The a priori tangent points and off-nadir angles reported by the instruments are determined differently: ACE-FTS [18] and MLS [22] provide geometric (*unrefracted*) values derived from the instrument pointing, without explicitly accounting for atmospheric refraction, whereas MIPAS [15] estimates these angles using the standard US76 atmospheric model, which includes refractive effects.

This choice directly affects the curvature of the line of sight, and consequently the altitude and location of the tangent point, whose prediction accuracy depends primarily on the assumed atmospheric model, as already shown in [24].

Once the atmospheric model is fixed, the corresponding off-nadir angles can be predetermined and used as input to value the accuracy of the tangent point prediction. To include the effects of the atmosphere in the ray-tracing computation, the atmospheric properties must first be discretized on a suitable grid, as detailed in Appendix A. The accuracy is then estimated as the displacement between the predicted tangent point and the actual tangent point computed using the ERA5 reanalysis data, taken as the reference *true* atmosphere. This displacement can be decomposed into a vertical shift in altitude, and an horizontal displacement along the angular level line. We focus on the tangent altitudes lower than 20 km, where the influence of the atmosphere is more significant and the curvature of the ray path is more pronounced [16]. Above this altitude, the effect of atmospheric refraction becomes negligible, as the lines of sight traverse less dense layers with smaller variations in the refractive index.

4 Results

In this section, we present two results obtained with the new ray-tracing algorithm. First, we evaluate the accuracy of tangent-points prediction depending on the Earth atmospheric model used to determine the off-nadir angles, following the methodology illustrated in Sect. 3 and using the CAIRT acquisition pattern as an example framework. Second, we demonstrate an independent application: the algorithm is used to estimate the MIPAS smoothing error resulting from the instrument's low horizontal resolution

4.1 Accuracy of the Tangent Point Prediction

In order to quantify the accuracy of the tangent points, we selected a test sun-synchronous orbit inclined at 98.7° to the equatorial plane, with an average altitude of 830 km, the ascending node at a latitude of 0°, and the UTC time of July 10, 2021, 12:00:00. We then applied the procedure of Sect. 3.1 to determine the off-nadir angles corresponding to the engineering tangent altitudes $\{5, 10, 15, 20\}$ km, using both the geometric and US76 atmosphere for the calculation. Then we performed the ray tracing using the *true* ERA5 atmosphere and measured the displacement of the tangent point. In the left panel of Fig. 2, we show the tangent altitudes obtained for the geometric model (dashed lines) and the US76 model (solid lines). These values are plotted as a function of the satellite polar coordinate in the orbital plane, which indicates the satellite's position along its circular orbit. A tolerance band of 500 m is also shown; this value corresponds to the goal of the *absolute geometric knowledge* as specified in the CAIRT Mission Required Document [10]. In the right panel, we show a zoom of the tangent altitudes predicted using the US76 model, in order to appreciate the variations.

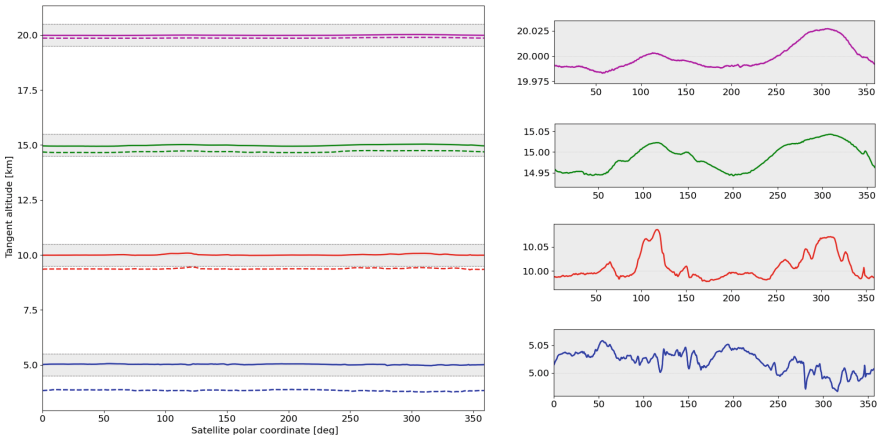


Fig. 2. Variation of the tangent altitude as a function of the satellite polar coordinate for a test orbit. Left panel: tangent altitudes obtained with the reference atmosphere along the orbit when the off-nadir angles are estimated with the geometric model (dashed line) and with the US76 model (solid line). A tolerance of 500 m is also shown. Right panel: zoom of the US76 plots

As shown in the figure, and as expected, the predictions from the geometric model are less accurate than those made with the US76 model. The off-nadir angles required to reach the engineering tangent altitudes are underestimated in the geometric model, as the curvature effect of the atmosphere is not accounted for. Consequently, the tangent point altitudes exhibit a negative bias. In detail, a line of sight starting from off-nadir angles that do not account for atmospheric refraction, reach tangent heights that, under realistic atmospheric conditions, are displaced by approximately 50 m at 20 km, increasing up to 1 km at 5 km relative to the predicted values.

These results indicate that the approximation introduced by the geometric model is too coarse, so we examine the US76 model. In Fig. 3 we show the vertical displacement using an extended range of engineering talgent altitudes from 5 to 40 km, with a 1 km step. As expected, the differences statistically increase when the engineering tangent altitude decrease, because the line of sight samples the lower and denser atmospheric layers. There are some exceptions due to the local atmospheric composition, for instance the maximal displacement of approximately 100 m is reached for the 10 km engineering tangent altitude.

The displacement of the tangent points can be measured not only through the vertical difference, as shown so far, but also through the polar coordinate difference, which provides complementary spatial information. Since the ray tracing is confined to the orbital plane, we can use the two dimensional coordinate system using φ, its angle along the Earth's elliptical surface, and z, its altitude above the Earth WGS84 surface, see Appendix A for more details. The polar displacement is then computed at a fixed engineering altitude as the length of

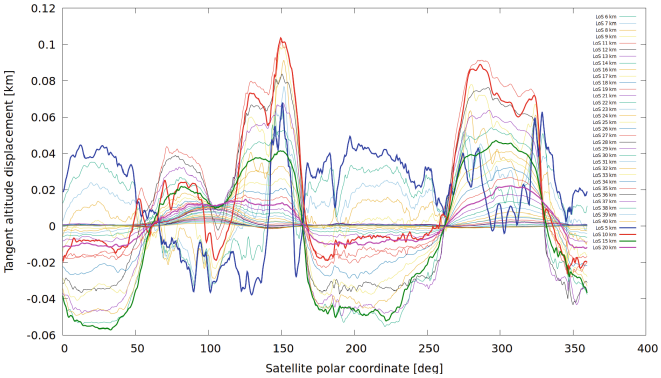


Fig. 3. Tangent altitude vertical displacement as a function of the polar coordinate of the satellite. The tangent altitudes considered in the previous figure are highlighted with a thicker line, and the color is maintained.

the arc between the angle φ of the engineering tangent point and that of the tangent point obtained with the ERA5 atmosphere.

We selected a sample of 14 consecutive orbits at the same UTC times as in the previous case, covering different latitude bands to provide a representative sampling of the whole Earth. Just as in the previous case, we focused on the $\{5, 10, 15, 20\}$ km engineering tangent altitudes. We report in Fig. 4 the scatter plots of the vertical and polar tangent point displacements between the ERA5 and the US76 tangent points, with a color gradient representing the position along the orbit.

As shown in the figure, the maximal polar displacement is about 1 km for the 20 km engineering tangent altitudes, and reaches 6 km for the lowest point. In particular, we note a positive correlation between vertical and polar displacements. This fact can be explained in the following way. When the vertical displacement is negative, the ERA5 tangent point altitude is lower than that of the US76 model. Consequently, the ERA5 line of sight must be longer to reach this lower altitude. A longer line of sight implies that the tangent point has a larger orbital shift with respect to the satellite. Since the satellite observes backward along its orbit, this corresponds to a smaller orbital coordinate and therefore a negative polar displacement. In addition, we observe a correlation also between the orbital coordinate and the displacement, independently of the starting longitude of the orbit, as it is visible in the color gradient of the points in the plots of Fig. 4. This suggests a relationship with the latitude bands, and consequently an influence of the climatic conditions.

In order to further investigate this correlation an additional test was performed comparing two orbits covering the same longitudinal band: one acquired on 10 July 2021 at 12:00:00 UTC, and the other on 10 January 2021 at 12:00:00 UTC. We focus on summer and winter, as the seasonal effect is maximized during these periods. We selected a fixed tangent altitude of 5 km as a represen-

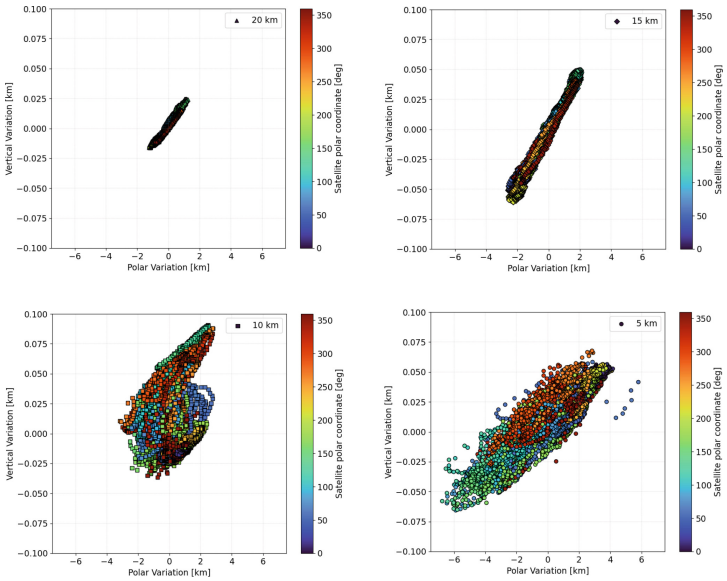


Fig. 4. Scatter plots illustrates the vertical and polar(horizontal) variation of the tangent point from off-nadir angles prediction, calculated as the difference between ERA5 and the US76 atmospheric model, as a function of the satellite's polar angle along the orbit. The plots includes data for four reference altitudes: 5 km, 10 km, 15 km, and 20 km, across 14 orbits.

tative case, since we expect more significant differences from the lowest tangent point. The difference in these two orbits is that the January orbit samples winter conditions in the Northern hemisphere and summer conditions in the Southern hemisphere. The seasonal conditions are reversed for the July orbit.

The results are shown in Fig. 5. The central panel shows the scatterplot of polar and vertical displacements between ERA5 and US76 model. Here, there are no significant differences in this plot. The auxiliary panels display the vertical (left) and polar (bottom) displacements as functions of the satellite's orbital coordinate. Note that, since the instrument looks backward, there is a negative shift of about 25° between the satellite and the tangent point orbital coordinate. Dashed line shows the approximate coordinate when the tangent point crosses the equator, highlighting the orbital intervals where the tangent point samples each hemisphere. As it can be seen there is a correlation between the hemisphere and the sign of both the vertical and polar displacements, confirming our previous assumption. Furthermore, the seasonal effect results in variations of about 5 km in the polar displacement, and of about 50 m in the altitude displacement, highlighting another impact of realistic atmospheric conditions on tangent point localization.

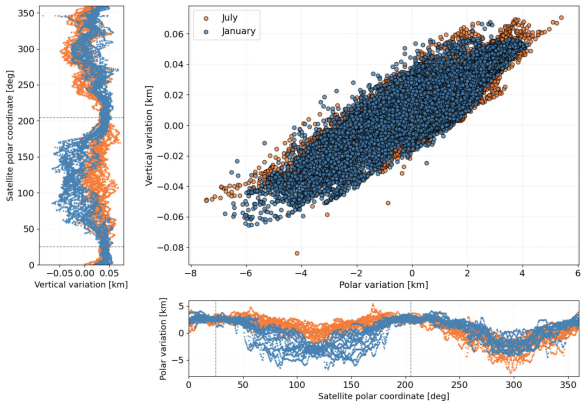


Fig. 5. Scatter plot for the 5 km tangent altitude. The plot shows the seasonal difference for one orbit in July (orange dots) and one orbit in January (blue dots) compares two distinct atmospheric conditions. The differences between the two atmospheric conditions are shown in terms of the vertical and polar variation in the tangent point. Dashed lines at $\sim 25^\circ$ and 205° indicate the approximate tangent point positions near the equator.

4.2 MIPAS Smoothing Error Estimation

The MIPAS Level-2 operational processor *Optimized Retrieval Model* (ORM) [25], is a forward and inversion code that in the last version [8] includes a dedicated ray-tracing module designed for the limb-viewing geometry. A direct comparison between the ORM ray-tracing algorithm and the new algorithm developed for CAIRT is not straightforward, due to fundamental differences in their design.

The ORM employs a different Earth geometry, defined as locally spherical and constructed around the tangent points, which are expressed in fixed polar coordinates. These coordinates are determined a priori by a Level-1 ray-tracing algorithm based on the US76 atmosphere. Each limb scan—defined by a sequence of tangent altitudes along the observation direction—is processed independently within this local spherical framework, where the local Earth radius is calculated from the tangent-point latitude.

More specifically, in the first step of the retrieval, the tangent-point temperatures and pressures for each scan sweep are derived, and the corresponding tangent altitudes are iteratively updated using hydrostatic equilibrium. The ORM ray tracing starts from the tangent point and proceeds bidirectionally up to the top of the atmosphere. In its latest version, ORM also accounts for horizontal gradients to correct biases introduced by the different atmospheric conditions sampled during the ascending and descending parts of the orbit [14].

The ray-tracing algorithm developed for CAIRT instead, as described in Sect. 3, starts from the satellite position and determines the atmospheric entry point along a straight line defined by the given off-nadir angle. The ray path

is then computed by numerically solving the Eikonal Eq. (1) until the line of sight exits the atmosphere. The selected atmospheric model, intersected by the line of sight, is discretized on a curvilinear coordinate system defined by altitude and polar angle along the WGS84 ellipsoid. Additional information regarding the numerical implementation, methodological improvements, and grid geometry can be found in the technical Appendices.

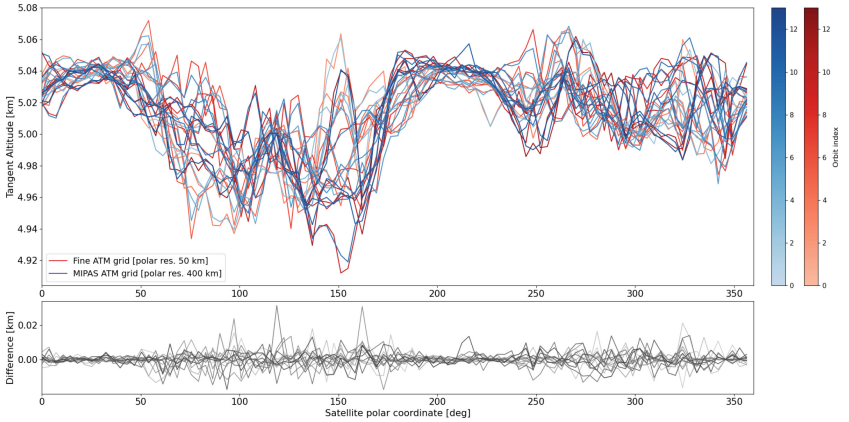


Fig. 6. The plot shows the variation of the tangent point altitude as a function of the satellite's polar coordinate, for a reference tangent altitude of 5 km. The blue curves correspond to values computed using the MIPAS horizontal resolution, while the red curves represent results obtained with a finer atmospheric grid. Tangent heights were evaluated across 14 consecutive satellite orbits, as indicated by the color gradients in blue and red. The subplot presents the residuals, defined as the differences between the tangent altitude estimates derived from the two spatial resolutions.

Our purpose is to quantify the smoothing error introduced by the coarser horizontal resolution of MIPAS acquisitions (about 400 km) compared to the finer sampling achievable with CAIRT (about 50 km). Because of the different approaches in ray tracing, a direct comparison using MIPAS retrieval code is not easy to implement. Thus, we have adopted the following methodology: we discretized the atmosphere using a horizontal grid of 50 km along the WGS84 ellipse, which we assume as the reference *true* atmosphere. This was done by defining the spacing between consecutive polar-angle grid points along the ellipse that form the grid. We then degraded the grid to 400 km, consistent with the horizontal resolution of MIPAS acquisitions. While the grid points are sampled using the same ERA5 data, bilinear interpolation is used to obtain the atmospheric properties needed to compute the refractive index at points within each grid cell. In the coarser resolution, linear interpolation is applied along the polar-angle direction between grid points spaced 400 km, while the interpolation in the altitude direction remains unchanged. This approach is consistent with the Level-2 MIPAS data and reflects how horizontal inhomogeneities are treated in

the ORM. We then compared the ray tracing, using the same MIPAS scan pattern, between the coarse and the fine atmospheric discretization. We assume that the bias due to the coarser resolution is an estimation of the MIPAS smoothing error.

To ensure global coverage, we again examined the set of 14 consecutive orbits already introduced in Subsect. 3.1. The results are presented in Fig. 6. Tangent altitudes computed using the atmosphere sampled at CAIRT resolution are presented in red, while those using the coarser MIPAS resolution are presented in blue. The lower panel shows the differences between the two. We observe that the smoothing error introduced by the MIPAS resolution remains globally below 25 m, an uncertainty directly related to its limited horizontal resolution. This error is much lower than the global uncertainty found in the identification of the tangent altitudes for MIPAS that can reach 200 m [24].

5 Conclusions

Accurate ray tracing is essential in limb-scanning interferometers to account for atmospheric refraction and to precisely determine the tangent point positions. In this paper, we upgraded the MIPAS ray-tracing algorithm, introducing new techniques to solve stability problems and improvements in geometric modeling. We provide the improvement introduced in the algorithm as a reference and use it for two specific analyses.

The first application is on a more accurate study on the relationship between off-nadir angles and tangent points altitudes. The off-nadir angles must be determined a-priori, using either a geometrical model that neglects atmospheric refraction or a standard atmospheric model such as US76. We first show that using the US76 model the predicted and true location of the tangent point are much closer than using the purely geometrical approach, as expected.

Our results show that using off-nadir angles derived from a standard atmosphere leads to tangent point displacements that increase at lower altitudes, reaching up to approximately 100 m in vertical displacement and up to 6 km in the horizontal displacement (i.e., along the polar coordinate of the atmospheric grid), while seasonal variability can introduce additional shifts of up to 50 m vertically and up to 1 km horizontally. This highlights the significant impact of realistic atmospheric conditions on tangent point localization, particularly in the lower troposphere.

In the second application, we estimated the smoothing error due to the coarse 400 km horizontal resolution of MIPAS. The smoothing error was quantified as the displacement of the tangent point when using an atmospheric horizontal grid of 50 km, considered the true state, compared to a 400 km grid, corresponding to the MIPAS resolution. We verified that the altitude error remains globally below 25 m, a value that is lower than the general tangent point error due to the atmosphere of about 200 m as estimated in [24].

Overall, as expected, these results quantify more accurately how using models as close as possible to the true atmospheric state reduces the error in tangent point localization for limb-instruments, particularly in the lower troposphere.

We used the CAIRT mission concept as a case study to explore limb-sounding scenarios where such challenges are expected. Although CAIRT was not selected as an operational mission, the ray-tracing remains relevant for any future limb-scanning missions.

As a final remark, we provide some additional considerations, also from a technical perspective, regarding the computation time.

The time required to compute a single LoS depends primarily on the length of its intersection with the atmosphere. Lower tangent altitudes result in longer atmospheric paths and consequently higher computational costs.

Although the computation time per LoS remains moderate on standard hardware, the large number of LoS paths required by satellite as CAIRT may lead to a computational bottleneck. As a reference, on a standard Intel i7 processor, the ray-tracing code requires approximately 2 h to process a full orbit consisting of 800 acquisitions, each containing 85 LoS.

Further improvements of the ray-tracing algorithm could be envisaged. We foresee two possible approaches:

(i) **Parallelism:** The calculation of each line of sight is independent of the others. A straightforward parallelization of the code should not require any modification to the underlying algorithm, and can be used to reduce the computational time.
(ii) **AI-based methods:** We have shown that off-nadir angles prediction depends also on seasonal variations. To improve the accuracy, this correlation should be incorporated in the atmospheric models, by selecting a climatological atmospheric state that varies with geolocation and time. Machine learning methods trained on a representative dataset can then be used to speed up the computation.

An additional optimization could involve reducing the number of lines of sight explicitly computed by the ray-tracing algorithm. Lines of sight corresponding to nearby engineering tangent altitudes are often strongly correlated, as they traverse nearly identical portions of the atmosphere. Therefore, it may be feasible to interpolate or emulate intermediate lines of sight using machine learning techniques trained on a reduced set of fully computed paths.

However, a key challenge arises from the level of detail required by the forward model: the relevant information is not limited to scalar quantities, but includes the precise geometry and trajectory of each ray path. This geometric fidelity must be preserved to ensure the accuracy of the radiative transfer calculations, making interpolation or emulation non-trivial.

Acknowledgments. The work was conducted within the framework of the CAIRT Analysis and Synergy with IASI-NG (CASIA) project of the Italian Space Agency (ASI), CUP: F93C23000430001.

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

CRediT Authorship Contribution Statement. Francesco Pio De Cosmo: Formal Analysis, Software, Testing, Writing – original draft, Writing – review & editing. **Luca Sgheri:** Methodology, Formal Analysis, Software, Writing – original draft, Writing – review & editing, Funding acquisition.

A Earth and LoS Geometry

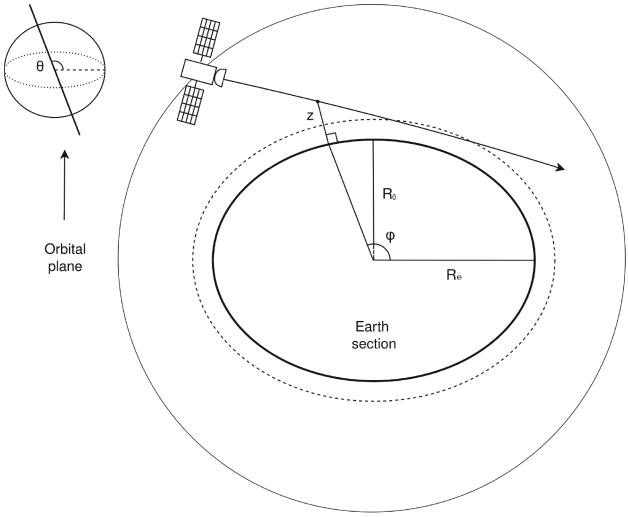


Fig. 7. Observation geometry for a satellite performing limb-viewing in a θ-inclined orbit, illustrating the effect of Earth's oblateness. The figure shows a section of the Earth (WGS84 model) in the satellite's orbital plane. The circle represents the satellite orbit, and the ellipse represents the intersection of the Earth with the orbital plane, with R_e and R_θ denoting the semi-major and semi-minor axes, respectively. The angle φ parameterizes the position of a point on the Earth's surface, and z is its altitude above the surface, measured normal to the ellipse. This allows the points along the line of sight to be expressed in the curvilinear coordinate system (φ, z). Orbit radius and Earth's oblateness are exaggerated for visual clarity.

Earth is represented according to the WGS84 model [19], a rotational ellipsoid with an equatorial (semi-major) radius of $R_e = 6378.137$ km and a polar (semi-minor) radius of $R_p = 6356.752$ km. This choice is essential to avoid large geometric approximations.

As discussed in the main text, the satellite is assumed to move along a circular, sun-synchronous orbit, allowing it to pass over the same Earth location at approximately the same local solar time. Such an orbit defines a plane inclined by an angle θ with respect to the equatorial plane, intersecting the Earth's ellipsoid as illustrated in Fig. 7.

If we restrict our analysis to the orbital plane with inclination θ relative to the equator, the problem can be simplified from three to two dimensions by considering the intersection between the Earth's ellipsoid and the orbital plane. This intersection forms an ellipse whose semi-major axis coincides with the equatorial radius R_e, while the semi-minor axis R_θ depends on the orbital inclination and can be calculated as

$$R_\theta = \sqrt{\frac{R_e^2 R_p^2}{R_p^2 \cos^2\theta + R_e^2 \sin^2\theta}}. \tag{7}$$

We note that $R_\theta = R_e$ for $\theta = 0$ (equatorial direction) and $R_\theta = R_p$ for $\theta = \pi/2$ (polar direction). Within the orbital plane, Earth's surface can be represented by an ellipse centered at Earth's center. In Cartesian coordinates, points on the ellipse can be expressed as $(x(\varphi), y(\varphi)) = (R_e \cos\varphi,\ R_\theta \sin\varphi)$, with $\varphi \in [0, 2\pi]$ the polar parametric angle of Earth's surface. Each point along this ellipse can be mapped to geodetic latitude and longitude, which allows the identification of the sub-satellite position. Atmospheric properties can then be extracted at these locations for a specific UTC.

The computation of the lines of sight extends from the satellite position up to the exit point from the atmosphere. The detectors of course capture the radiation traveling along the line of sight in the opposite direction, reaching the instrument after sampling the atmosphere. Points along these rays, initially calculated from the Eikonal equation in Cartesian coordinates, are then mapped to a curvilinear coordinate system (φ, z), where φ is the polar angle of Earth's surface in the orbital plane and z represents the altitude above the surface, as shown in Fig. 7.

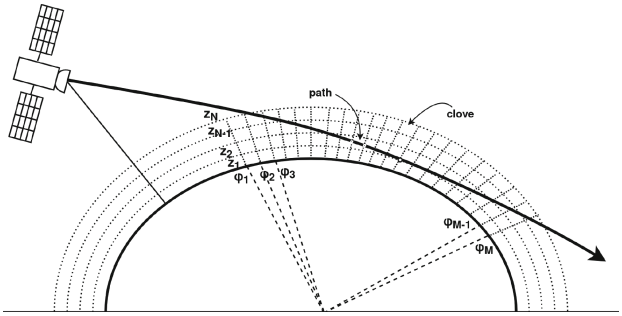


Fig. 8. Atmospheric grid used in the ray-tracing model. The atmosphere is discretized into altitude levels $z_1, \ldots, z_N$ and angular coordinates $\varphi_1, \ldots, \varphi_M$, forming a 2D grid. Each grid cell, called a *clove*, represents a sector of the discretized atmosphere. The segment of a LoS that crosses a clove defines a *path*.

Atmospheric properties are first extracted at sub-satellite geolocations. They are then interpolated into altitude levels $z_1, \ldots, z_N$ and angular coordinates $\varphi_1, \ldots, \varphi_M$, forming a 2D grid $\{\varphi_1, \ldots, \varphi_M;\ z_1, \ldots, z_N\}$, as illustrated in Fig. 8.

This grid defines the curvilinear discretization used in the ray-tracing calculations. A *clove* is defined as a sector of an elliptical annulus, determined by the atmospheric grid nodes, while a *path* corresponds to the intersection of a LoS with a clove. Within each clove, atmospheric properties at intermediate points are obtained by bilinear interpolation of the four surrounding grid points, except for the vertical interpolation in pressure, which follows an exponential rule. These interpolated properties are then used to determine the refractive index, which governs the local curvature of the line of sight. This approach ensures that the discretization captures both horizontal and vertical variations of the atmosphere along the ray path, while maintaining consistency with the underlying curvilinear coordinate system (φ, z).

References

1. Bohren, C.F., Huffman, D.R.: Absorption and Scattering of Light by Small Particles. John Wiley & Sons, New York (1983)
2. Boone, C.D., Bernath, P.F., Walker, K.A., McLeod, S.D.: Ace-fts version 3.6 retrieval updates. J. Quant. Spectroscopy Radiative Transfer **247**, 106939 (2020). https://doi.org/10.1016/j.jqsrt.2020.106939
3. Born, M., et al.: Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light, 7 edn. Cambridge University Press (1999)
4. Buehler, S.A., Eriksson, P.: Microwave limb sounding of the troposphere. In: Platt, U., Stutz, J. (eds.) The Remote Sensing of the Earth's Atmosphere and Surface. University of Bremen Press (2018). https://www.iup.uni-bremen.de/materials/remsensingbook/04_Remote_Sensing_Microwaves.pdf
5. Burrows, J., Hölzle, E., Goede, A., Visser, H., Fricke, W.: Sciamachy–scanning imaging absorption spectrometer for atmospheric chartography. Acta Astronaut. **35**(7), 445–451 (1995). https://doi.org/10.1016/0094-5765(94)00278-T
6. Ciddor, P.E.: Refractive index of air: new equations for the visible and near infrared. Appl. Opt. **35**, 1566–1573 (1996). https://doi.org/10.1364/AO.35.001566
7. Copernicus Atmosphere Monitoring Service (CAMS): Cams global greenhouse gas reanalysis (egg4). https://atmosphere.copernicus.eu/cams-global-greenhouse-gas-reanalysis (2020)
8. Dinelli, M.B., et al.: The esa mipas/envisat level2-v8 dataset: 10 years of measurements retrieved with ORM v8.22. Atmospheric Measure. Techn. **14**(12), 7975–7998 (2021). https://doi.org/10.5194/amt-14-7975-2021
9. Edlén, B.: The refractive index of air. Metrologia **2**, 71–80 (1966). https://doi.org/10.1088/0026-1394/2/2/001
10. ESA: Report for mission selection: Earth explorer 11 candidate mission cairt, european space agency, noordwijk, the netherlands, esa-eopsm-cair-rp-4797, 230 pp (2025). https://doi.org/10.5281/zenodo.15606819
11. European Space Agency: Polar and sun-synchronous orbit (2020). https://www.esa.int/ESA_Multimedia/Images/2020/03/Polar_and_Sun-synchronous_orbit
12. Fischer, H., et al.: Mipas: an instrument for atmospheric and climate research. Atmos. Chem. Phys. **8**(8), 2151–2188 (2008). https://doi.org/10.5194/acp-8-2151-2008
13. Hersbach, H., et al.: The era5 global reanalysis. Q. J. R. Meteorol. Soc. **146**(730), 1999–2049 (2020). https://doi.org/10.1002/qj.3803

14. Kiefer, M., et al.: Impact of temperature field inhomogeneities on the retrieval of atmospheric species from MIPAS IR limb emission spectra. Atmospheric Measure. Techn. **3**, 1487–1507 (2010). https://doi.org/10.5194/amt-3-1487-2010
15. Kiefer, M., et al.: Characterization of mipas elevation pointing. Atmos. Chem. Phys. **7**(5), 1615–1625 (2007). https://doi.org/10.5194/acp-7-1615-2007
16. Livesey, N.: Limb sounding, atmospheric. In: Encyclopedia of Remote Sensing, pp. 344–348. Springer, New York, NY (2014).https://doi.org/10.1007/978-0-387-36699-9_87
17. Mathar, R.J.: Calculated refractivity of water vapor and moist air in the atmospheric window at 10 μm. Appl. Opt. **43**, 928–932 (2004). https://doi.org/10.1364/AO.43.000928
18. NASA Atmospheric Science Data Center: Ace-fts ascii data usage and file format for v3.5 (2016). https://avdc.gsfc.nasa.gov/PDF/ACE-SOC-0026-1B-ACE-FTS_ascii_data_usage_and_fileformat_for_v3.5-15March2016.pdf
19. National Imagery and Mapping Agency: Wgs 84: The world geodetic system 1984 (2000). https://www.ngs.noaa.gov/WGS84/
20. NOAA, NASA, U.S. Air Force: U.S. Standard Atmosphere, 1976. U.S. Government Printing Office, Washington, D.C. (1976). https://ntrs.nasa.gov/citations/19770009539
21. Raspollini, P., et al.: Level 2 processor and auxiliary data for ESA version 8 final full mission analysis of MIPAS measurements on Envisat. Atmospheric Measur. Techn. **15**(6), 1871–1901 (2022). https://doi.org/10.5194/amt-15-1871-2022
22. Read, W.G., Shippony, Z., Snyder, W.V.: Earth observing system (EOS) microwave limb sounder (MLS) forward model: algorithm theoretical basis document. Technical Report JPL D-18130/CL#04-2238, NASA, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA (2004)
23. Remedios, J.J., et al.: MIPAs reference atmospheres and comparisons to v4.61/v4.62 MIPAs level 2 geophysical data sets. Atmospheric Chemis. Phys. Discuss. **7**, 9973–10017 (2007).https://doi.org/10.5194/acpd-7-9973-2007
24. Ridolfi, M., Sgheri, L.: Characterization of model errors in the calculation of tangent heights for atmospheric infrared limb measurements. Atmospheric Measur. Techn. **7**(12), 4117–4122 (2014). https://doi.org/10.5194/amt-7-4117-2014
25. Ridolfi, M., et al.: Optimized forward model and retrieval scheme for mipas near-real-time data processing. Appl. Opt. **39**(8), 1323–1340 (2000). https://doi.org/10.1364/ao.39.001323
26. Riese, M., et al.: Earth explorer 11 mission candidate cairt: phase 0 system study overview. In: 2023 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), p. 10641615 (2023). https://doi.org/10.1109/IGARSS47720.2023.10641615
27. University of Washington: Strive. https://atmos.uw.edu/strive/ (2025)
28. Waters, J., et al.: The earth observing system microwave limb sounder (EOS MLS) on the aura satellite: a 17-year record of temperature, ozone, and other atmospheric constituents. Atmospheric Measur. Techn. **15**(6), 1583–1610 (2022). https://doi.org/10.5194/amt-15-1583-2022

Short Papers

Artificial Intelligence for Photovoltaic Power Forecasting in Energy Management

Guillermo Blanco[2], Marco Melgarejo[1(✉)], Alejandro Pérez[1], and Camilo Palazuelos[2]

[1] Centro Tecnológico CTC, Parque Científico y Tecnológico de Cantabria, 39011 Santander, Spain
mmelgarejo@centrotecnologicoctc.com

[2] Universidad de Cantabria Av. de los Castros s, 39005 Santander, n, Spain ,

Abstract. Photovoltaic power forecasting plays a key role in the efficient management of renewable energy systems, particularly in installations operating under zero-injection constraints. This paper presents the development of a forecasting model trained on data collected from a pilot photovoltaic installation located in Santander, Spain. The proposed approach combines historical power generation data with meteorological variables to predict the expected generated power and support more efficient energy management decisions. The paper describes the main stages of the workflow, including data preprocessing, feature engineering, model training, and evaluation. The resulting model achieves an R^2 score of 0.92, showing strong predictive performance and demonstrating the potential of machine learning methods for photovoltaic energy forecasting.

Keywords: Machine Learning · Photovoltaic · Artificial Intelligence · Photovoltaic Forecasting · Time Series Prediction · Renewable Energy · XGBoost · Energy Optimization

1 Introduction

The Spanish photovoltaic market is divided into producers and self-consumers. Producers are those who sell their energy to the electrical grid. In contrast, self-consumers do not sell their production; instead, they use it for their own consumption. In Spain, regulatory constraints have been introduced for certain photovoltaic installations. Some self-consumers, typically those with an installed capacity above 100 kW, are required to comply with the zero-injection requirement, ensuring that no power is fed into the grid.

The zero-injection requirement is a regulation that prohibits the injection of surplus photovoltaic generation into the electrical grid. Therefore, self-consumers must adjust their photovoltaic production so that it does not exceed the electrical demand of the installation [8].

Energy flexibility in this context refers to the ability to adapt electricity demand according to external factors such as tariffs or regulatory constraints.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 173–185, 2027.
https://doi.org/10.1007/978-3-032-30539-8_11

Previous studies have shown that predictive control of photovoltaic production can reduce total energy costs by around 15–25%, while also helping to reduce CO_2 emissions. According to the International Energy Agency, in 2025 solar photovoltaic energy made a significant contribution to reducing global emissions, avoiding on the order of 1.5 Gt of CO_2 anually [4].

Energy optimization can be divided into two types according to the time horizon considered. Short-term optimization focuses on hourly energy production in order to minimize costs and increase efficiency over short periods of time. Long-term optimization aims to determine the optimal dimensions of the installation and its components while minimizing installation costs. To estimate the electrical demand of an installation, multiple time-dependent factors must be considered, as shown in Eq. (1):

$$D(t) = B(t) + P(t) + N(t) + A \tag{1}$$

where $B(t)$ represents the basic operating cycle, $P(t)$ the peak demands that introduce maximum-demand events without overlap, $N(t)$ the stochastic noise fluctuations, and A the baseline consumption, ensuring realistic conditions [3].

Spain increased its installed photovoltaic capacity from 11 GW in 2020 to more than 32 GW by the end of 2024 [11]. As a result, photovoltaic power has become one of the country's main electricity generation technologies, even surpassing wind power. At present, photovoltaic energy accounts for 17% of total national electricity generation, making Spain one of the leading countries in solar energy in Europe.

At the international level, photovoltaic energy represents 46% of renewable capacity additions, and the International Energy Agency (IEA) estimates around 3.6 TW of new renewable energy capacity before 2030, of which approximately 80% is expected to come from photovoltaic sources. As 2030 approaches, the photovoltaic industry is expected to strengthen its role as the leading renewable energy technology [4].

Compared with other renewable energy sources, photovoltaic power generally shows a favorable environmental profile. Emissions associated with its manufacturing process are lower than those of other energy sources, such as wind power.

Despite these achievements, the photovoltaic industry still faces important challenges, one of the main ones being the price of electricity. As this technological sector has expanded rapidly, internal competition has increased significantly, leading to a decline in the selling price of solar energy. This situation has directly affected the profitability of photovoltaic installations [6]. If corrective measures are not taken soon, the situation is expected to worsen by the end of 2026.

The goal of this work is to develop a prediction model that primarily helps installations operating under the zero-injection rule to manage their demand through forecasting. In addition, the proposed model can also be useful for intelligent energy management systems in general. Its purpose is to predict the expected power generated by the photovoltaic installation.

2 Related work

Photovoltaic power forecasting has gained prominence amid Spain's rapid solar expansion. Early models employed statistical data but limited by non-linear weather dependencies.

Machine learning advancements shifted toward tree-based ensembles. *Random forests* and *Extra Trees* improved handling of meteorological inputs like irradiance and temperature, achieving robust short-term predictions on UCI datasets. *Gradient boosting*, particularly *XGBoost*, outperformed these via optimized regularization and feature importance, with R^2 scores exceeding 0.90 in benchmarks [12].

As shown in Table 1, in Spain tools such as Shirenda [15] leverage machine learning on REE data (2015–2020) for nationwide yield estimation, supporting zero-injection compliance under RD 244/2019. Recent studies forecast rooftop PV with tuned *XGBoost* models and emphasize real-time AEMET integration, but they still overlook edge deployment [5].

These contributions enable demand optimization, saving 15-25% costs, but gaps persist in low-resource hardware forecasting and hybrid *skforecast* [13] with *XGBoost* for real pilot data [2], motivating our methodology.

Table 1. Comparison of related forecasting studies

Work	Method	Horizon	Key Metric	Limitation
Chen (2016) [5]	*XGBoost*	Short-term	R^2: 0.92	Offline focus
Zhang (2023) [9]	*LSTM-AdaBoost*	Short	RMSE: 0.043	Synthetic data
Shirenda_PV (2025) [10]	ML on REE	Long-term	N/A (yield est.)	No real-time

This work further investigates the application of machine learning models, especially *XGBoost*, to photovoltaic power forecasting. In particular, it analyzes how these models capture non-linear relationships between meteorological variables and photovoltaic generation, thereby improving forecasting accuracy and supporting more informed operational decisions in solar installations.

3 Methodology

3.1 Data source

The data were collected from a real photovoltaic installation located on the roof of CTC in Santander, Spain, over the course of one year. The power generation and meteorological datasets were merged and temporally synchronized. The same procedure was also applied to another photovoltaic installation in Lorca, Murcia, Spain. However, the present study focuses primarily on the Santander case, while the Lorca installation is used only as a complementary example in the practical applications section.

3.2 Dataset preprocessing

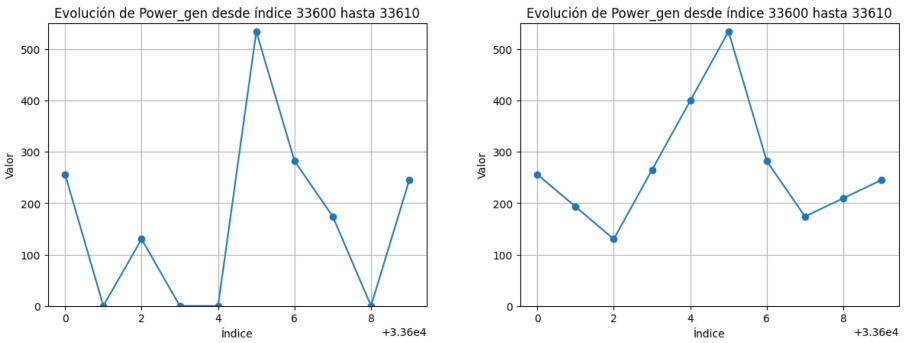
When developing a forecasting model, the first step is to ensure that the dataset used for training is properly prepared and processed. The data must reflect the conditions of the target installation so that the model can learn representative local patterns. This has a direct impact on forecasting accuracy.

Once the dataset was assembled, an additional analysis was performed to detect and remove outliers that could negatively affect model accuracy.

The 25th percentile is the value below which 25% of the observations fall, while the 75th percentile is the value below which 75% of the observations fall. The interquartile range (IQR) is defined as the difference between these two percentiles [14]. Outlier removal results in a moderate reduction in the dataset size, which is acceptable, although it also creates gaps in the time series.

Some daytime samples recorded between sunrise and sunset showed zero power generation values. These values were considered physically implausible in the present modelling stage, since zero-injection operation was not included during model training. Instead, the zero-injection constraint is addressed in a subsequent optimization layer, further discussed in Section 4.3, where equipment demand can be scheduled according to the forecasted photovoltaic production in order to avoid surplus injection into the grid.

Consequently, these anomalous zero-valued samples were attributed to possible data logging errors and corrected by linear interpolation. Retaining them would have introduced artificial discontinuities in the generation signal and could have increased the prediction error. Fig. 1 illustrates this preprocessing step: the left panel shows the original signal with anomalous zero-valued samples, whereas the right panel shows the corrected signal after interpolation.



(a) Before interpolation (b) After interpolation

Fig. 1. Effect of linear interpolation on the time series. The left panel shows the original signal with anomalous zero-valued samples, whereas the right panel shows the corrected signal after interpolation.

Finally, some rows contained NaN values in one or more columns, and these rows were removed. After completing the preprocessing stage, the final dataset was obtained, as summarized in Table 2.

Table 2. Dataset descriptive statistics

Statistic	Power_gen(W)	temperature(C^{O})	windSpeed(km/h)	solarRad(W/m^2)	moist(%)	rain(mm)	pressure(hPa)
count	41821.00	41821.00	41821.00	41821.00	41821.00	41821.00	41821.00
mean	86.63	16.32	6.87	101.45	79.29	0.53	30.13
std	156.52	4.22	5.99	162.07	10.71	1.24	0.20
min	0.00	4.44	0.00	0.00	46.00	0.00	29.60
25%	0.00	12.90	3.22	0.00	73.00	0.00	30.00
50%	0.00	16.30	4.83	0.00	82.00	0.00	30.20
75%	94.00	19.80	9.66	153.00	88.00	0.00	30.30
max	780.00	30.60	27.40	631.00	96.00	5.40	30.70

Since the variables in the dataset are measured on different scales, feature scaling was applied. This step must be performed only after splitting the data into training and test subsets; otherwise, information from the training data would leak into the test set. This issue, known as data leakage, can lead to overly optimistic performance estimates.

Scaling maps the variables to a common range, typically between 0 and 1. In this work, this was performed using *MinMaxScaler*.

After preprocessing, the model development and training stage was carried out.

3.3 Model development

Time series are data ordered chronologically. Forecasting consists of predicting future values of a time series by exploiting its past behavior in an autoregressive manner, optionally combined with external variables through tools such as *skforecast* [2]. In this work, exogenous variables such as solar irradiance, temperature, and wind speed were used together with endogenous features derived from previous observations, including mean, median, maximum, and minimum values. Unlike standard regression settings, time plays a central role in time-series forecasting; however, the problem can still be addressed using regression-based models such as *Gradient Boosting* or *Random Forest*.

A forecasting model based on *XGBoost* was trained. The optimal hyperparameters were selected by means of a *GridSearchCV* procedure. After training, feature importance analysis was used to assess the relative contribution of each predictor to the model output, as shown in Fig. 2.

As expected in a tree-based time-series forecasting model, the most influential predictors are the nearest lagged values, namely *Lag_1* and *Lag_2*, which represent the observations recorded 5 and 10 minutes earlier. However, this should not overshadow the relevance of the exogenous variables.

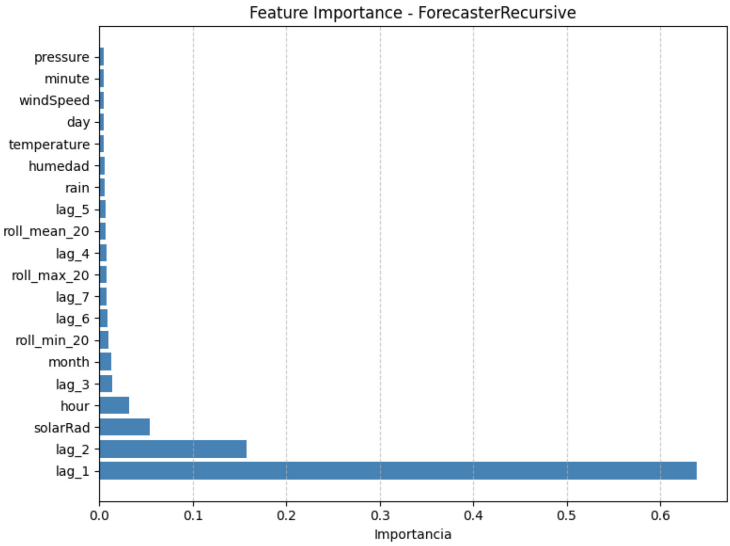


Fig. 2. Feature importance.

In particular, solar irradiance ranks as the third most important feature overall and the most important non-autoregressive predictor, which is consistent with the physical behavior of photovoltaic generation. The hour of the day also appears among the most relevant variables, further supporting the temporal and physical consistency of the model.

In addition, summary statistics derived from previous samples, such as rolling minima and maxima, are also present in the top-ranked features, indicating that local temporal context contributes useful complementary information. Therefore, although the lagged variables dominate the ranking, relying only on *Lag_ 1* and *Lag_ 2* would be an oversimplification, as the remaining predictors help reduce forecasting error and improve the robustness of the model.

It is also essential to verify that the model is not overfitting the training data. Overfitting occurs when the model memorizes the training set instead of learning generalizable predictive patterns, which typically leads to excellent performance on seen data but poor results on unseen data. A common way to assess this behavior is by comparing the evolution of the training and validation errors across boosting iterations. If the validation error increases substantially while the training error continues to decrease, the model can be considered overfitted. Figure 3 shows that the training and validation curves evolve consistently, indicating that the model does not exhibit significant overfitting.

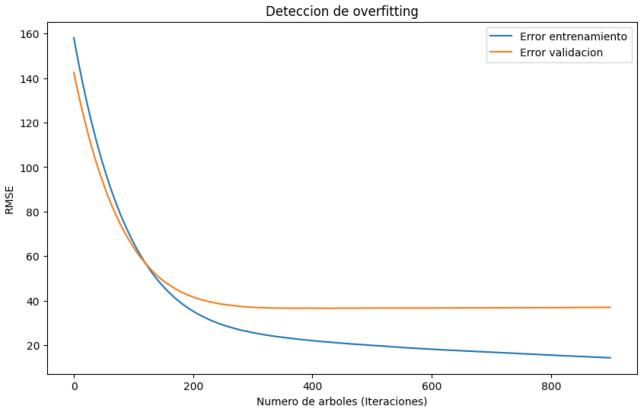


Fig. 3. Training and validation error across iterations.

4 Results

4.1 Prediction metrics

After generating predictions on the test set, the performance metrics reported in Table 3 were obtained.

Table 3. Prediction metrics

MAE	16.52 W
RMSE	40.39 W
R^2	0.92

The results indicate that the proposed model achieves a satisfactory level of predictive accuracy on the test set. This is especially relevant considering that the case study corresponds to a photovoltaic installation located in Cantabria, where frequent meteorological fluctuations and variable cloud cover lead to rapid changes in solar irradiance. Therefore, the forecasting task is more challenging than in scenarios with more stable weather conditions.

In particular, the R^2 value shows that the model explains a large proportion of the variance in the target variable, while the MAE and RMSE values indicate that the prediction error remains moderate in relation to the typical range of photovoltaic power generation observed in the dataset. Overall, these results suggest that the model is robust and suitable for short-term photovoltaic forecasting in a real operating scenario with significant weather variability.

4.2 Graphical representation

A useful way to assess the model performance is to compare the predicted values with the observed ones. Fig. 4 presents this comparison over the full test set, where the predictions are shown in red and the actual values in blue.

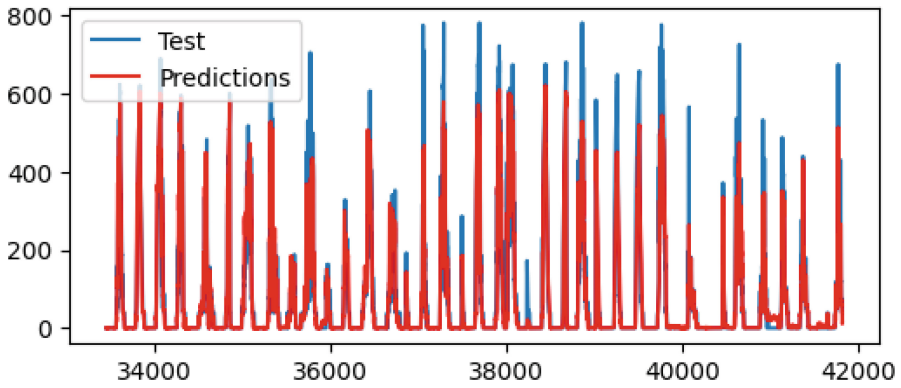


Fig. 4. Prediction through the whole dataset.

A more detailed evaluation can be performed over shorter forecasting horizons, such as the next 24 or 48 hours. Fig. 5 shows the 48-hour forecasting example, again using red for predicted values and blue for observed values. The similarity between both curves indicates that the model is able to capture the overall trend of the power generated by the installation.

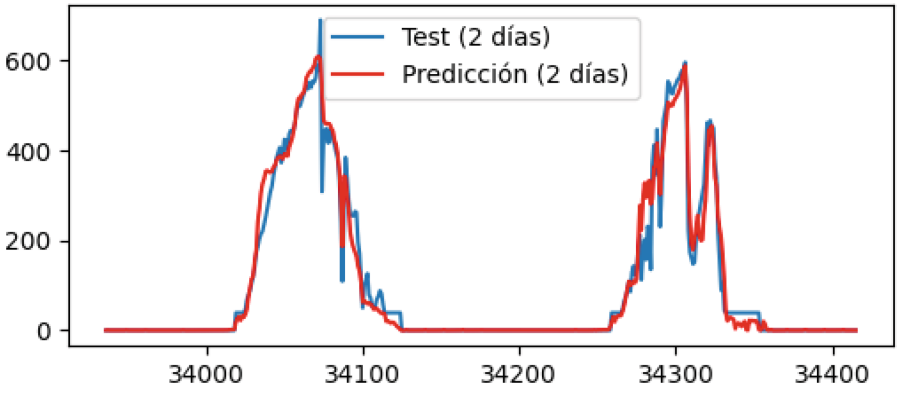


Fig. 5. Prediction of the next 48 hours.

Another way to evaluate model performance is to compare predicted and observed values in a scatter plot rather than along the time axis. Fig. 6 presents

this comparison, where the closer the points are to the dashed red diagonal, the better the agreement between predictions and observations. The predicted values are shown on the Y-axis and the observed values on the X-axis.

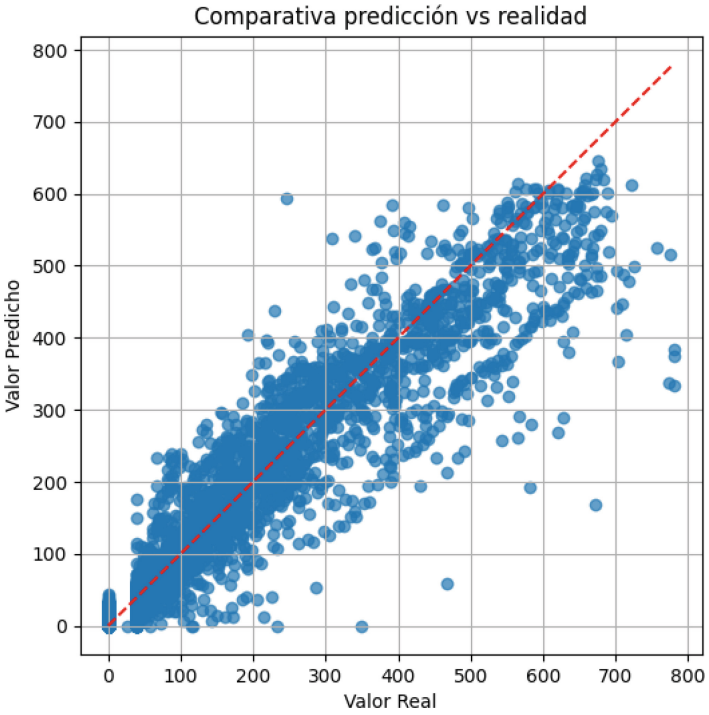


Fig. 6. Predicted vs. real values for the full test set.

4.3 Practical applications

Although the main case study of this paper is the Santander installation, the Lorca installation is included as a complementary example to illustrate the practical value of the proposed forecasting approach for industrial and commercial energy management.

From an operational perspective, the main usefulness of the forecasting model lies not only in point predictions, but also in deriving aggregated indicators of photovoltaic availability over decision-relevant time windows. In practice, daily forecasts can be summarized as average production levels for morning and afternoon shifts, as shown in Fig. 7. These indicators provide a simple representation of the photovoltaic energy expected to be available in each shift, together with the overall daily average.

Such information can be directly used to support load-balancing strategies. In industrial environments, several loads may be flexible, including pumping systems, compressed-air equipment, thermal processes, charging systems, or non-critical production tasks. If higher photovoltaic generation is forecast for a given shift, these loads can be shifted or intensified during that period to increase self-consumption. In this way, the forecast becomes a decision-support tool for adapting demand to the expected renewable generation profile.

This application is especially relevant under zero-injection constraints. In a strict zero-feed-in regime, all photovoltaic energy must be consumed locally at each time step; otherwise, export-limiting systems enforce curtailment or disconnection of the photovoltaic plant. Therefore, relying only on anti-export hardware may ensure compliance, but it can also waste usable renewable energy and increase implementation costs. By contrast, forecast-based demand adaptation allows a larger share of the locally generated energy to be used before curtailment becomes necessary. Both strategies can therefore be combined: export-limiting hardware can remain as a safety layer, while predictive energy management acts as the main mechanism to maximize photovoltaic utilization.

Once these forecasted averages are available, different automatic optimization strategies can be applied to schedule flexible loads. A first option is a heuristic or metaheuristic scheduler that assigns shiftable loads to the periods with the highest predicted photovoltaic availability. Previous works [1,7] have explored particle swarm optimization and other multi-objective approaches to improve self-consumption and reduce grid dependence. Heuristic multi-agent coordination schemes have also been proposed for scheduling shiftable devices, offering high-quality solutions with lower computational cost. In addition, robust multi-objective formulations can explicitly account for photovoltaic uncertainty while optimizing economic and environmental criteria. All these approaches are compatible with the present application, where the goal is to synchronize flexible demand with forecasted photovoltaic production.

Forecast uncertainty must also be considered, since it increases as the prediction horizon moves further away from the present time. For this reason, indicators such as those in Fig. 7 can be complemented with error bars or confidence intervals, allowing the operator or the energy management system to distinguish between more and less reliable forecast horizons. This enables a hierarchical planning strategy: day-ahead forecasts can define an initial schedule, while intraday updates progressively refine it as new meteorological information and more accurate short-term predictions become available. In other words, the load-balancing plan can be continuously corrected in a rolling-horizon manner as the target operating period approaches.

Therefore, the practical contribution of the proposed forecasting model goes beyond estimating photovoltaic generation. Its main added value lies in enabling predictive load scheduling and energy optimization strategies that increase self-consumption, reduce unnecessary curtailment, and mitigate the operational drawbacks associated with zero-injection requirements.

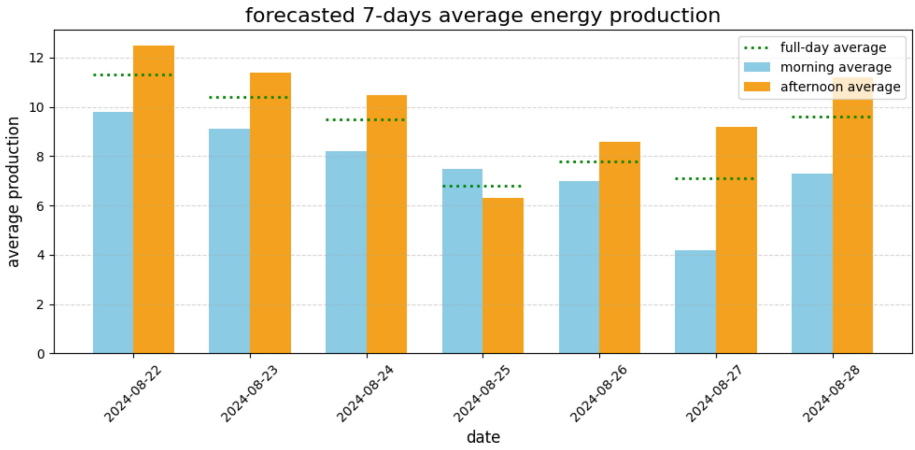


Fig. 7. Average production profile for the Lorca installation, included as a complementary application example.

5 Conclusions and future steps

The development of a forecasting model is only meaningful when it serves a practical purpose. As stated earlier, the main objective of this work was to support the electrical management of an installation operating under the zero-injection rule. Thanks to the proposed model, it is possible to estimate the amount of energy that will be generated and adjust the electrical demand accordingly.

Compared to the state of the art, the proposed methodology demonstrates competitive performance while addressing relevant gaps in the literature. The model achieved an R^2 of approximately 0.92, which is in line with previously reported short-term forecasting results. More importantly, whereas several recent studies rely on synthetic data or do not consider real-time edge deployment, the proposed approach combines skforecast and XGBoost on real pilot data. This strengthens the practical relevance of the methodology and supports its potential for low-resource hardware forecasting.

As future work, the model could be extended by incorporating external meteorological forecasts, such as those provided by AEMET, as exogenous variables during prediction, rather than relying only on historical local measurements from the on-site weather station. Although these forecasts are usually generated on spatial grids, they can be adapted to the exact installation location through statistical downscaling or simpler geographical interpolation methods.

In addition, the same approach could provide access to a broader historical meteorological record, enabling the construction of a more representative training dataset. This would improve the robustness of the model and further support its deployment in practical photovoltaic energy management applications.

Data Availibility. The data supporting the results of this study are available from the corresponding author upon request.

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

Funding. The presented work is part of the project titled *Posicionamiento Estratégico de Especialización Inteligente y Sostenible para las Factorías del FUTuro de CANtabria (FUTCAN)*, which is being carried out at the Centro Tecnológico CTC. This entity has received, for the FUTCAN project, co-financed assistance from the European Regional Development Fund (ERDF) through the 2021–2027 Cantabria Operational Program, under the funding line *Grants for research projects with high industrial potential of excellence technology agents for industrial competitiveness (TCNIC).*











References

1. Abdalla, M.A.A., Min, W., Bing, W., Ishag, A.M., Saleh, B.: Double-layer home energy management strategy for increasing pv self-consumption and cost reduction through appliances scheduling, ev, and storage. Energy Rep. **10**, 3494–3518 (2023). https://doi.org/10.1016/j.egyr.2023.10.019
2. Amat Rodrigo, J., Escobar Ortiz, J.: skforecast (Sep 2023). https://doi.org/10.5281/ZENODO.8382788, https://doi.org/10.5281/ZENODO.8382788, accessed on: May 8, 2026
3. Castán-Lascorz, M.A., Mselle, B.D., Zabaleta, L.L., León, M., Arroyo, J.: Energy flexibility optimization in industry: A hybrid approach with synthetic data evaluation. In: 2025 IEEE International Conference on Engineering, Technology, and Innovation (ICE/ITMC). pp. 1–10. Valencia, Spain (Jun 2025). https://doi.org/10.1109/ICE/ITMC65658.2025.11106527
4. IEA: Global energy review 2026: Co2 emissions. Tech. rep., International Energy Agency, Paris (2026), https://www.iea.org/reports/global-energy-review-2026/co2-emissions, accessed on: May 8, 2026. Licence: CC BY 4.0
5. IJPEDS: Article download 24057/15126 (2026), https://ijpeds.iaescore.com/index.php/IJPEDS/article/download/24057/15126, accessed on: March 6, 2026
6. International Renewable Energy Agency: Renewable power generation costs in 2023. Tech. rep., International Renewable Energy Agency, Abu Dhabi (Sep 2024), https://www.irena.org/Publications/2024/Sep/Renewable-Power-Generation-Costs-in-2023, accessed on: May 6, 2026
7. Luz, G.P., Brito, M., Sousa, J., Vieira, S.: Coordinating shiftable loads for collective photovoltaic self-consumption: A multi-agent approach. Energy **229**, 120573 (2021). https://doi.org/10.1016/j.energy.2021.120573, https://www.sciencedirect.com/science/article/pii/S0360544221008227
8. Ministerio para la Transición Ecológica: Real decreto 244/2019, de 5 de abril, por el que se regulan las condiciones administrativas, técnicas y económicas del autoconsumo de energía eléctrica. BOE-A-2019-5089, pp. 35674–35719 (2019), https://www.boe.es/eli/es/rd/2019/04/05/244, accessed on: February 23, 2026

9. PMC: Article pmc11002275 (2026), https://pmc.ncbi.nlm.nih.gov/articles/PMC11002275/, accessed on: March 6, 2026
10. PV Magazine: Nueva base de datos abierta para estimar la producción fotovoltaica en españa (2025), https://www.pv-magazine.es/2025/02/03/nueva-base-de-datos-abierta-para-estimar-la-produccion-fotovoltaica-en-espana/, accessed on: March 6, 2026
11. Red Eléctrica de España: Informe de energías renovables: potencia instalada, solar fotovoltaica (2025), https://www.sistemaelectrico-ree.es/es/informe-de-energias-renovables/sol/potencia-instalada/solar-fotovoltaica-solpotencia, information prepared using provisional data up to January 2026. Accessed on: May 6, 2026
12. ScienceDirect: Article s2352484725001490 (2026), https://www.sciencedirect.com/science/article/pii/S2352484725001490, accessed on: March 6, 2026
13. skforecast: skforecast documentation (v0.20.1) (2026), https://skforecast.org/0.20.1/index.html, accessed on: March 7, 2026
14. Statistics How To: Interquartile range (iqr): What it is and how to find it (2025), https://www.statisticshowto.com/probability-and-statistics/interquartile-range/, accessed on: December 1, 2025
15. Sánchez-Hernández, G., Jiménez-Garrote, A., López-Cuesta, M., Galván, I.M., Aler, R., Pozo-Vázquez, D.: A novel method for modeling renewable power production using era5: Spanish solar pv energy. Renewable Energy **240**, 122120 (2025). https://doi.org/10.1016/j.renene.2024.122120, https://doi.org/10.1016/j.renene.2024.122120

Economic Assessment of Inclusive Tourism in a Silk Road Heritage Destination: A Monte Carlo Simulation Study of Samarkand, Uzbekistan

Bahodirhon Safarov[1,2], Habibullo Hasanov[2(✉)], Makhina Buzrukova[1], and Elena Salnikova[2]

[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

khabibulloeco@gmail.com, h.hasanov@univ-silkroad.uz

Abstract. This study assesses the economic potential of inclusive tourism in Samarkand as a Silk Road heritage destination. The analysis applies Monte Carlo simulation to evaluate two key aspects: the profitability of inclusive services at the hotel level and the expenditure structure of an inclusive tourist during a five-day trip. The hotel model shows that the most favorable service price is around 140,000 UZS, where expected profit reaches its highest level and the probability of obtaining a positive result remains relatively strong. The results also indicate that prices above 230,000 UZS significantly reduce the likelihood of profitability. From the consumer perspective, the expenditure model reveals that inclusive tourism generates considerable spending not only on accommodation but also on transport, recreation, shopping, and related support services. Among all expenditure components, caregiver service has the strongest effect on total expenditure, followed by recreation, shopping, and transport. These findings suggest that inclusive tourism in Samarkand should be viewed not only as a social necessity but also as a promising economic market segment. The study contributes to the literature by offering a simulation-based approach to destination-level economic planning for inclusive tourism.

Keywords: Inclusive Tourism · Accessible Tourism · Monte Carlo Simulation · Tourism Expenditure · Economic Assessment

1 Introduction

Inclusive tourism is now considered not only a social but also an important economic aspect of tourism. According to the World Health Organization, about 1.3 billion people in the world, or about 16% of the global population, live with a significant degree of disability [1, 2]. This reinforces the need to develop tourism services in a more accessible and safe way for all.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 186–196, 2027.
https://doi.org/10.1007/978-3-032-30539-8_12

According to the UN Tourism and ISO 21902 approach, accessible tourism for all involves ensuring equal and convenient access at all levels of the tourism value chain - information, transport, accommodation, services and use of tourist facilities [3, 4]. Therefore, inclusive tourism is not a separate social service, but a direction that enhances the destination's competitiveness, the scope of services, and the possibilities for sustainable development.

This issue is especially relevant for Silk Road heritage destinations. Historical cities such as Samarkand are distinguished by their cultural heritage, international tourist flows and the diversity of services [5–10]. However, the economic aspects of inclusive tourism in such destinations, in particular the composition of tourist spending, the acceptable price of services and the probabilistic assessment of market potential, have not yet been sufficiently studied. Most existing research focuses on barriers, infrastructure, and ease of use, while economic modeling is less common [11–14].

In this context, the purpose of this study is to assess the economic impact of inclusive tourism in the Silk Road heritage destination of Samarkand. The article analyzes the costs of inclusive tourism, the acceptable range of service prices, and economic efficiency for tourism entities using a Monte Carlo simulation. Such an approach enhances decision-making under uncertainty and helps to justify inclusive tourism as a factor of sustainable economic development. Monte Carlo simulation has been used in tourism economics as a useful tool for assessing uncertainty and economic impact [15, 16].

2 Literature Review

In the existing scientific literature, inclusive tourism has been studied mainly in relation to reducing physical, informational, and organizational barriers to the use of tourism services. At the same time, some studies view accessibility as a factor that enhances a destination's competitiveness; in particular, Porto et al. showed that a region's competitive advantage increases with tourism accessibility [17, 18]. European-wide assessments have revealed that the accessible tourism segment has significant economic potential: in 2012, demand from travelers with special access needs generated economic benefits of hundreds of billions of euros in the European Union, and this demand is projected to increase further [19, 20]. These conclusions indicate the need to view inclusive tourism not only as a mechanism for social support but also as a source of economic growth and the expansion of the service market [21].

In tourism economics research, tourist spending has emerged as a distinct field, and Brida and Scuderi have shown that analyzing the volume of spending and the factors that affect it at the microeconomic level is important for assessing the real economic impact of tourism [22]. Subsequent studies have found that tourist spending is significantly influenced by trip duration, travelers' income, group composition, mode of transport, and accommodation choice [23]. It has also been noted that the itemized expenditure approach, that is, the study of expenses by accommodation, food, transport and other components, provides a deeper economic conclusion than the total amount of expenditure [24]. However, most of the work in this area has focused on the flow of ordinary tourists, and the cost structure of inclusive tourists who use additional services, such as accompaniment, adapted transport, or special accommodation, has not been sufficiently

covered. This gap is especially noticeable in Silk Road destinations such as Samarkand, where heritage and services are combined [19].

The issue of assessing the tourism market under uncertainty is also gaining increasing attention in the literature. Recent work has shown that Monte Carlo simulation is a useful tool for assessing tourism demand, economic impact, and possible scenarios [15, 25]. In particular, this method is significant in that it shows a range of possible outcomes, not just the average value. However, the existing literature still has very limited application of Monte Carlo simulation to the assessment of inclusive tourism costs, the acceptable range of service prices, and economic efficiency [26]. Therefore, assessing the economic potential of inclusive tourism through a simulation approach in the case of Samarkand helps fill the existing scientific gap in this topic.

3 Materials and Methods

This study used the Monte Carlo simulation method to assess the economic potential of inclusive tourism in the Samarkand region. The analysis was conducted in two directions. In the first direction, the profitability of a hotel providing inclusive services was modeled. In the second direction, the composition of the 5-day travel expenses of an inclusive tourist and their possible changes were estimated. The model was used in conjunction with a scenario analysis approach suitable for decision-making under uncertainty.

The number of inclusive tourists visiting the hotel in the model was expressed based on the Binomial distribution:

$$P(N = k) = \binom{n}{k} p^k (1 - p)^{n-k} \tag{1}$$

Here, denotes the number of tourists, the number of trials, and the probability of attracting a tourist. In the model, the number of potential customers was set at 5,000, while the probability of acceptance was assumed to be 0.55. Market prices for inclusive services were modeled using a Normal distribution:

$$P_m \sim N\left(\mu, \sigma^2\right) \tag{2}$$

where = 150,000 UZS represents the average market price and = 60,000 UZS is the standard deviation. It was assumed that when the hotel's offered price was lower than the market price, the service would be purchased; otherwise, demand would decline and the risk of loss would increase. Total hotel costs were calculated as the sum of fixed cost and variable cost:

$$TC = FC + VC \tag{3}$$

where = 36,000,000 UZS, while was specified using a Normal distribution as (100,000,000, 30,000,000^2). Hotel profit was defined as follows:

$$\pi = R - TC \tag{4}$$

where denotes revenue and total cost. The probability of making a profit was calculated as the share of simulation runs in which profit was positive. In total, 1,329 simulation outcomes were generated for the hotel model.

The expenses for an inclusive tourist on a 5-day trip were divided into two categories: basic and contingent. Basic expenses included transportation, accommodation, food, shopping, and recreation. The minimum, expected, and maximum values for these expenses were determined based on the questionnaire results and data from the Booking.com platform. These parameters are presented in Table 1.

Table 1. Main Expenditure Items of an Inclusive Tourist for a 5-Day Trip, USD.

Expenditure item	Minimum value	Expected value	Maximum value
Transport	50	150	800
Accommodation	50	100	500
Food	50	125	600
Shopping	25	250	600
Recreation	100	250	800

The PERT distribution was used for the main costs because it produces a relatively realistic distribution based on minimum, most likely, and maximum values. The expected value in the PERT distribution is calculated as follows:

$$E(X) = \frac{a + 4m + b}{6} \tag{5}$$

where is the minimum value, is the most likely value, and is the maximum value. This approach assumes that costs tend to be shaped by the most likely value.

Probable costs included medical services and caregiver services. Their probability of occurrence was also estimated based on the questionnaire and is presented in Table 2. These costs were modeled using a combination of the Bernoulli and PERT distributions.

Table 2. Probabilistic Expenditure Items of an Inclusive Tourist, USD.

Expenditure item	Probability of occurrence	Minimum value	Expected value	Maximum value
Medical service	15%	50	200	1000
Caregiver service	25%	250	500	1000

The likely costs were estimated as follows:

$$B \sim \text{Bernoulli}(p) \tag{6}$$

where takes the value 0 or 1, i.e., it represents the cases where the cost did not occur or did occur. The total cost is then expressed as:

$$C_{total} = \sum_{i=1}^{5} \mu_i + \sum_{j=1}^{2} e_j \tag{7}$$

where is the main cost, is the potential cost. The calculations were performed in MS Excel, and 1500 random values were generated to assess the distribution of total costs, the cumulative distribution, and the relationship between cost types and total cost.

4 Results

The results of the Monte Carlo simulation showed that pricing is a crucial issue for a hotel providing inclusive tourism services in Samarkand. Figure 1 shows the changes in expected profit and the probability of profit as a function of the service price. As shown in the graph, the hotel's highest expected profit occurs when the price of the inclusive service is set at 140 thousand UZS. At the same time, the probability of making a profit at this point is also quite high, and according to the model, the hotel achieves a positive result in about 60% of cases. This situation shows the importance of setting prices in the inclusive services market, not maximum prices, but in a competitive range close to the market price. Based on the results, the 80–150 thousand UZS range for the optimistic scenario, the 150–230 thousand UZS range for the baseline scenario, and the price range above 230 thousand UZS for the pessimistic scenario were considered typical. Especially at prices above 230,000 UZS, the fact that the profit probability drops to almost zero indicates that the hotel's risk of loss increases sharply.

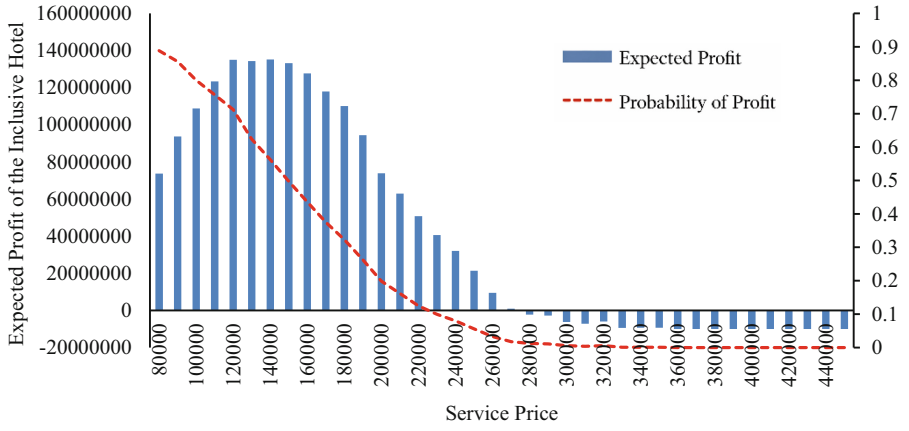


Fig. 1. Distribution of Expected Profit and Profit Probability.

The results of the simulation of the composition of expenses for an inclusive tourist on a 5-day trip are shown in Fig. 2. The graph shows the distribution of transportation, accommodation, food, shopping and recreation expenses for 1,500 simulations. Among these distributions, recreation expenditure is particularly distinguished by its relatively wide range of variations. In most cases, recreation expenses range from 200 to 400 US dollars, indicating that they account for a significant share of an inclusive tourist's total expenses. Although the remaining expense types also fluctuate within a certain range, their degree of change and their impact on total expenses differ. In this regard,

Fig. 2 shows not only the distribution of expenses by individual but also the services that constitute the structural composition of an inclusive tourist's expenses.

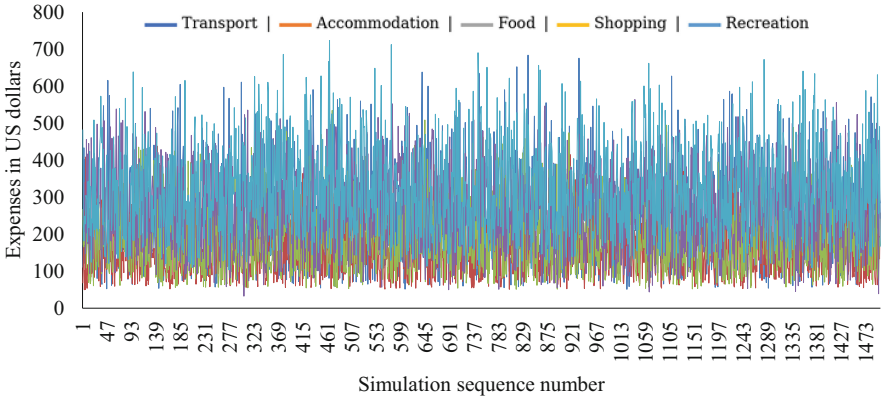


Fig. 2. Distribution of Main Expenditure Items across Simulations.

The cumulative distribution of the total 5-day expenses is presented in Fig. 3. This graph clearly shows that inclusive tourist expenses are not fixed at a single value but lie within a range. According to the model results, in 72% of cases, the total expenses are less than 1564 USD. In 50% of the experiments, the 5-day expenses are equal to or less than 1300 USD, and in 10% of the cases, they are 920 USD. These results show that consumer expenses can be quite high in the inclusive tourism segment. If these values are converted to daily terms, 1300 USD over 5 days averages 260 USD per day, and 920 USD averages approximately 184 USD per day. Based on this, the inclusive tourism market can be considered a segment with relatively high revenue potential for tourism entities.

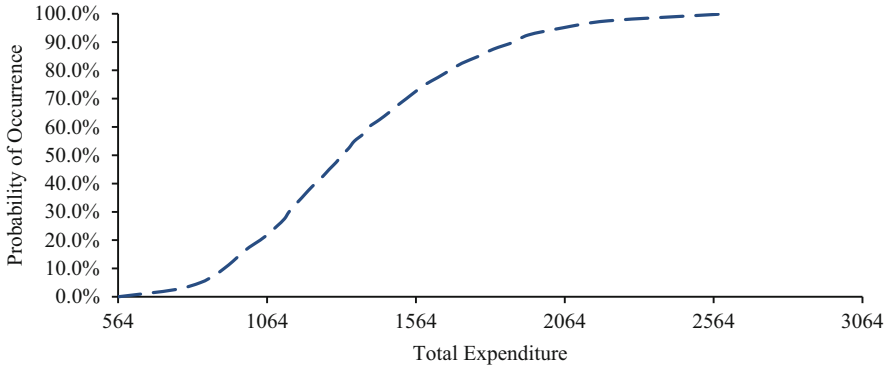


Fig. 3. Cumulative Distribution of Total Expenditure of an Inclusive Tourist.

The impact of expenditure types on total expenditure was assessed using correlation analysis, as shown in Fig. 4. The results indicate that caregiver services had the greatest impact, accounting for 67% of total expenditure. The following were recreation (35%), shopping (35%), transport (33%), medical service (28%), food (24%), and accommodation (23%). This result shows that additional and accompanying services are the most important economic factor in inclusive tourists' expenditure. In other words, as inclusive tourism develops, revenue is generated not only in the hotel and accommodation sector but also through related services such as caregiving, medical services, recreation, and trade. Therefore, Fig. 4 is one of the most important results showing that inclusive tourism has a multi-sectoral economic impact.

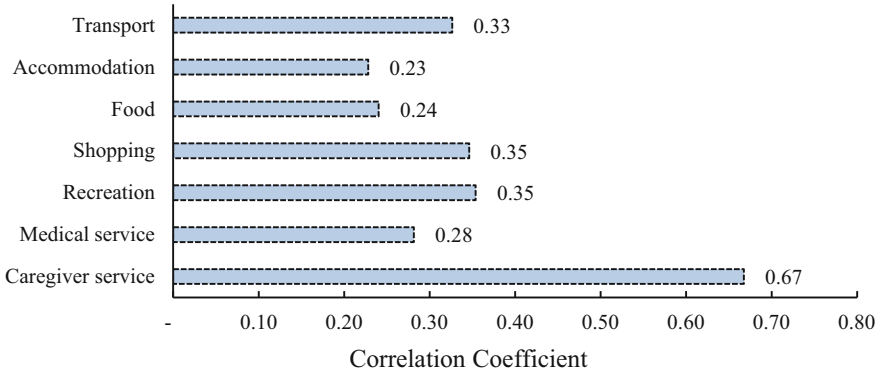


Fig. 4. Correlation between Expenditure Components and Total Expenditure.

In general, the results indicate that inclusive tourism is not only a socially important direction for the Samarkand region but also a promising market from an economic perspective [26]. At the hotel level, the right pricing strategy increases the likelihood of profitability, and at the consumer level, spending volume over 5 days confirms that this segment creates sufficient value for the services market. In particular, the strong influence of caregivers and related services indicates the need to expand these services to support the development of inclusive tourism in Samarkand.

5 Discussion

The results of this study show that inclusive tourism is a trend with not only social but also clear economic potential in the context of a Silk Road heritage destination like Samarkand. This conclusion is consistent with European studies showing that the economic contribution of the accessible tourism market is significant; in particular, people with access needs typically travel with an average of 1.9 companions, and the economic contribution increases further as accessibility improves [27].

The optimal service price, determined by the hotel model at around 140 thousand soums, and a sharp decrease in profit probability after 230 thousand soums, partially confirm and partially clarify some previous studies. For example, Liu found that travelers

with disabilities are more willing to pay for certain accessible travel products, especially accessible accommodation and equipped transport [28]. However, our results show that this willingness is constrained by the overall travel budget, as inclusive tourists also incur additional costs for accommodation, transport, recreation, medical support, and caregivers. Therefore, in the Samarkand context, it is more important to balance the pricing strategy with the service package than to simply increase prices. This is also consistent with hotel-level research showing that the economic impact of accessibility is more evident through social innovation and proper service organization [21].

Another important result is that the largest impact on total expenditure was from caregiver services. This finding is consistent with the companion effect in accessible tourism studies in Europe [27]. However, in the case of Samarkand, this effect was relatively stronger. This can be explained by the dispersed location of historical sites, the complexity of physical movement in some areas, and the higher need for additional organizational support along heritage routes. Thus, the inclusive tourism economy here is built on a tourist-plus-support system model rather than an individual-tourist model.

The significant role of recreation, shopping, and transportation expenses also underscores the need to assess expenses not as a single total, but by their components. This result is consistent with studies on itemized expenditure: Marrocu et al. showed that trip duration, income, group composition, and mode of transport significantly affect tourist expenditure, while Massidda et al. found that the determinants of the cost components differ [23, 24, 29]. Our results confirm this general pattern but show one important difference: while accommodation is often the central focus of ordinary tourists, in inclusive tourism, additional services, such as caregiver services and medical support, significantly alter the economic structure. In this regard, the development of inclusive tourism in Samarkand requires the formation of a single service package that includes caregiver support, medical support, adapted transport, and clear access information, not limited to physical accessibility [30]. Recent studies on the accessible tourism value chain have also identified information completeness and service interconnections as weak points [31, 32].

6 Conclusion

This study showed that inclusive tourism in Samarkand is not only socially significant but also a direction with clear economic potential. The results of the Monte Carlo simulation confirmed that the right pricing strategy is important in the inclusive services market, and at the hotel level, profitability is maintained within a certain price range. At the same time, an analysis of the 5-day expenses of inclusive tourists showed that this segment can also be a real source of income for accommodation, transport, recreation, and other related services.

One of the most important results of the study is that caregiver services and other additional services play a decisive role in the total costs. This means that, in the development of inclusive tourism, it is necessary to create a single package of services, not limited to physical accessibility, but also including caregiver support, medical support, adapted transport, and accurate access information. Overall, this study offers a useful approach for planning inclusive tourism on economic grounds and for making practical decisions in Silk Road heritage destinations such as Samarkand.

Funding. 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. Accessible Tourism. (2026). https://www.untourism.int/accessibility-tourism. Last Accessed 9 Mar 2026
2. Disability. (2023). https://www.who.int/news-room/fact-sheets/detail/disability-and-health. Last Accessed 9 Mar 2026
3. Tourism and related services — Accessible tourism for all — Requirements and recommendations. (2021). https://www.iso.org/standard/72126.html
4. World Tourism Organization, Fundación ONCE, Normalización Española (eds): How to Apply ISO Standard 21902: Accessible Tourism for All – Recommendations for Tour Operators, Travel Agencies and Travel Agents. UN Tourism (2024). https://doi.org/10.18111/9789284425860
5. Nurullayeva, N. et al.: ANALYSIS OF SAMARKAND'S REGISTAN SQUARE AS a MASTERPIECE OF ISLAMIC ARCHITECTURE afts. **32**, 113–119 (2025). https://doi.org/10.70102/afts.2025.1732.113
6. Jumanazarov, S., Kamilov, A., Kiatkawsin, K.: Impact of Samarkand's destination attributes on international tourists' revisit and word-of-mouth intention. Sustainability. **12**, 5154 (2020). https://doi.org/10.3390/su12125154
7. Safarov, B. et al.: Prospects of agrotourism development in the region. Economies. **12**, 321 (2024). https://doi.org/10.3390/economies12120321
8. Tuxtabekov, K., Buzrukova, M., Hasanov, H., Salnikova, E.: Forecasting tourism services exports for sustainable development: a case study of Uzbekistan. In: Gervasi, O., Murgante, B., Garau, C., Karaca, Y., Faginas Lago, M.N., Scorza, F., Braga, A.C. (eds.) Computational Science and its Applications–ICCSA 2025 Workshops, pp. 176–186. Springer Nature, Switzerland, Cham (2026). https://doi.org/10.1007/978-3-031-97663-6_15
9. Patterson, I., Aslam, M.U.: Agritourism as an emerging sustainable tourism industry in Uzbekistan. Sustainability. **16**, 7519 (2024). https://doi.org/10.3390/su16177519
10. 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
11. Celestino, S., Garofano, A., Masiello, B., Izzo, F., Bonetti, E.: Disability and marketing: a bibliometric analysis and systematic literature review. Ital. J. Mark. **2024**, 311–337 (2024). https://doi.org/10.1007/s43039-024-00098-3
12. Makuyana, T., Du Plessis, E.: Exploring the obstacles and opportunities related to disability-inclusive diversity management in the local tourism sector of South Africa. Tour. Manag. **111**, 105238 (2025). https://doi.org/10.1016/j.tourman.2025.105238
13. Rickly, J., Hansen, M.: Advancing accessible tourism geographies. Tour. Geogr. **27**, 1363–1376 (2025). https://doi.org/10.1080/14616688.2025.2593982
14. Torres-Castillo, Y., Ortiz-Pérez, O.L.: Accessibility in rural tourist destinations: an investigation on barriers and perspectives in communities of the Las Tunas tourist destination. Rev. interam. ambient. tur. **20**, 25–38 (2024). https://doi.org/10.4067/s0718-235x2024000100025
15. Salgado Barandela, J., Barajas, A., Sanchez-Fernandez, P.: Implications for tourism management of including uncertainty in the estimation of the economic impact of sports events. SBM. **13**, 181–194 (2023). https://doi.org/10.1108/SBM-11-2021-0138

16. Gupta, A. et al.: Economic impact of nature-based tourism. PLoS One. **18**, e0282912 (2023). https://doi.org/10.1371/journal.pone.0282912
17. Porto, N., Rucci, A.C., Ciaschi, M.: Tourism accessibility competitiveness. A regional approach for Latin American countries. Investig. Reg., 75–91 (2018)
18. 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
19. Belgium, G.K.: University of Surrey, NeumannConsult, ProAsolutions: Economic Impact and Travel Patterns of Accessible Tourism in Europe–Final Report: Summary. European Commission, DG Enterprise and Industry, Brussels (2014)
20. Rebstock, M.: Economic Benefits of Improved Accessibility to Transport Systems and the Role of Transport in Fostering Tourism for all. Organisation for Economic Co-operation and Development (OECD), International Transport Forum, Paris (2017)
21. López-Gamero, M.D., Pereira-Moliner, J., Pertusa-Ortega, E.M., Molina-Azorín, J.F., Tarí, J.J.: Accessibility, competitiveness and performance relationships: the mediating role of social innovation. Int. J. Hosp. Manag. **128**, 104185 (2025). https://doi.org/10.1016/j.ijhm.2025.104185
22. Brida, J.G., Scuderi, R.: Determinants of tourist expenditure: a review of microeconometric models. Tour. Manag. Perspect. **6**, 28–40 (2013). https://doi.org/10.1016/j.tmp.2012.10.006
23. Marrocu, E., Paci, R., Zara, A.: Micro-economic determinants of tourist expenditure: A quantile regression approach. Tour. Manag. **50**, 13–30 (2015). https://doi.org/10.1016/j.tourman.2015.01.006
24. Massidda, C., Piras, R., Seetaram, N.: Analysing the drivers of itemised tourism expenditure from the UK using survey data. Ann. Tour. Res. Empir. Insights. **3**, 100037 (2022). https://doi.org/10.1016/j.annale.2022.100037
25. Samadi Parviznejad, P., Bahrami, M.: Uncertainty analysis of tourism components in Tabriz. IJIMES. **1**, 1–14 (2021). https://doi.org/10.52547/ijimes.1.3.1
26. López-del-Pino, F., Grisolía, J.M., Ortúzar, J.D.D.: Understanding the elements of the tourism accessibility chain. Travel Behav. Soc. **42**, 101161 (2026). https://doi.org/10.1016/j.tbs.2025.101161
27. 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. Geojornal Tour. Geosites. **58**, 484–491 (2025). https://doi.org/10.30892/gtg.58145-1430
28. European Commission: Guidelines for the Implementation of a Customized Accessible Tourism Programme. (2016). https://ec.europa.eu/programmes/erasmus-plus/project-result-content/5769d3b8-0658-4003-8c3b-6309331269b5/guidelines_for_the_implementation_of_a_customezied_accessible_tourism_program.With_disclaimer.pdf
29. Lyu, S.O.: Which accessible travel products are people with disabilities willing to pay more? A choice experiment. Tour. Manag. **59**, 404–412 (2017). https://doi.org/10.1016/j.tourman.2016.09.002
30. Pai, W.-T., Phan, K.-T., Pham, C.-V., Lee, J.-M., Hsieh, C.-J.: Analyzing tourist expenditures incurred on long-, medium-, and short-haul trips to Taiwan. Ann. Tour. Res. Empir. Insights. **5**, 100138 (2024). https://doi.org/10.1016/j.annale.2024.100138
31. 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., Murgante, B., Garau, C., Karaca, Y., Faginas Lago, M.N., Scorza, F., Braga, A.C. (eds.) Computational Science and Its Applications–ICCSA 2025 Workshops, pp. 405–414. Springer Nature, Cham (2026)

32. Heyerdahl Nyquist, T.E., Jahnsen, R.B., Bliksvaer, T., Koncul, A., Antonsen, K.M.: Accessible tourism – nature-based activities for all, needs for assistive devices and information among people with disabilities. J. Outdoor Recreat. Tour. **52**, 100982 (2025). https://doi.org/10.1016/j.jort.2025.100982

On Battery Storage Management by Sparse Linear Programming

Ana Maria A. C. Rocha[1(✉)], Pablo Guerrero-García[2], and Eligius M. T. Hendrix[3]

[1] Centro ALGORITMI/LASI, Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Portugal
arocha@dps.uminho.pt
[2] Applied Mathematics,Universidad de Málaga, 29071 Málaga, Spain
[3] Computer Architecture, Universidad de Málaga, 29071 Málaga, Spain

Abstract. This paper formulates a problem of optimal battery storage management as a sparse linear programming (LP) problem. We focus on the potential of solving this problem by an augmented Lagrangian (AL) approach. The problem maximizes revenue exploiting volatility of the electricity price. The model deals with equalities to describe the energy balance and contains bounds that reflect the operational limits. We investigate an AL-based algorithm with a first stage for infeasibility detection and an iterative scheme for convergence to a basic feasible solution. Numerical experiments using real electricity price data from Energia XXI on four days (31st October, 1st November, 3rd November, and 5th November 2025) demonstrate fast convergence within 2 to 14 iterations. The daily net revenue for a home situation varies between 0.7690 € and 2.3186 € depending on price volatility.

Keywords: Linear Programming · Battery Storage Management · augmented Lagrangian

1 Introduction

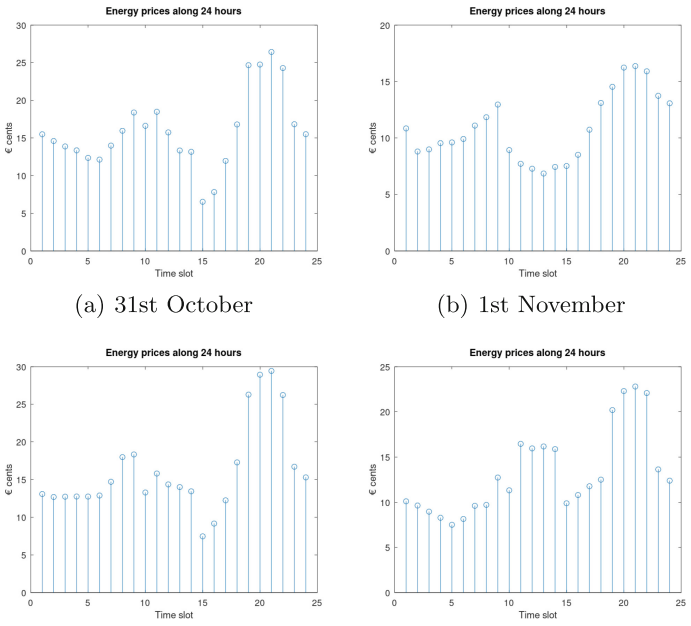
Battery energy storage systems (BESS) provide a solution to deal with variability in renewable energy production. BESS can be used to store excess energy generated during periods of high production and release it when generation is low. This leads to flexibility and stabilizing the provision. Similar observations have been discussed in recent reviews, where optimization of BESS has become central to the integration of renewables in smart grids [2,8].

This work has been supported by FCT Fundação para a Ciência e Tecnologia within the R&D Unit Project Scope UID/00319/2025 - Centro ALGORITMI (ALGORITMI/UM) https://doi.org/10.54499/UID/00319/2025 and a teacher mobility grant from Fundación Carolina and the Grupo Tordesillas 2025–2026.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 197–205, 2027.
https://doi.org/10.1007/978-3-032-30539-8_13

Variable electricity prices allow consumers to sell energy back to the grid when there is surplus and to buy and store when prices are low. Optimization-based strategies, including linear and stochastic formulations, have shown great potential to maximize economic benefits in such environments [1].

Currently, family households are able to buy a home battery, where we focus on a variant that can store 20 kWh. In contracts, the varying price is determined due to an auction system and is known one day in advance for each hour of the day. The dynamic electricity pricing and cyclic behaviour is sketched by the data provided in Fig. 1. This setting allows an individual consumer to buy during low price hours and sell during hours where the price is high. To optimize the revenue, this optimization problem can be captured by a Linear Programming (LP) problem with simple balance constraints and bounds.



(a) 31st October (b) 1st November

(c) 3rd November (d) 5th November

Fig. 1. Electricity prices for different days provided by `www.energiaxxi.com`.

The resulting problem is sparse. Recently, specific LP solution methods have been developed to deal with such sparse problems [5]. One of our questions is to adapt this approach to deal with upper bounds in an augmented Lagrangian setting. Our research investigates the potential of sparse LP formulations to efficiently solve the underlying battery storage optimization problem. Specifically, we focus on the potential of recent advances in sparse solvers and augmented Lagrangian (AL) methods [5,6].

To investigate the question, we introduce the mathematical optimization model in Sect. 2 and place it in the context of literature. We analyze the sparsity of the underlying model in Sect. 3 and show how modern LP solvers can exploit the structure in Sect. 4. Section 5 summarizes our findings.

2 Mathematical Description and Literature Embedding

The operation of management storage systems provides an advantage in revenue in systems where we are dealing with photovoltaic generation systems [7]. Other works have introduced linear and non-ideal battery models that better capture charge/discharge efficiencies and degradation effects without excessive computational burden [4]. The result of recent development in the energy conversion is that governments implement and promote systems in which the consumer is dealing with a varying price during the day in order to stabilize demand.

We describe an underlying LP problem in (1)-(6) that includes the basics of revenue optimization of a home battery for an individual consumer dealing with a price P_t that varies each hour t of the day and the possibility to return energy to the grid at the same price with an additional cost C.

The variables involved in the model are s_t, that is the amount stored in battery at the end of slot t, b_t is the quantity delivered back to grid in slot t and q_t refers to the quantity taken from grid in slot t. Table 1 resumes the notation and used symbols of the optimization problem.

In the model, (3) represents a balance equation and (2) ensures that the stored capacity is the same at the end and the beginning of the day. Similar formulations with extensions were used in previous studies of off-grid and grid-connected systems [1,3].

$$\max_{s,b,q} \sum_{t=1}^{T} P_t(b_t - q_t) - C \sum_{t=1}^{T} b_t \tag{1}$$

$$\text{s.t.} \quad s_0 = s_{T-1} - b_T + q_T, \tag{2}$$

$$s_t = s_{t-1} - b_t + q_t, \qquad t = 1, \ldots, T-1 \tag{3}$$

$$s_t \in [0, \overline{s}] \tag{4}$$

$$b_t \in [0, \overline{b}] \tag{5}$$

$$q_t \in [0, \overline{q}] \tag{6}$$

The sparse linear programming formulation for battery storage management will be tested using real-time electricity prices obtained from Energia XXI provided by `www.energiaxxi.com` for four days (31st October, 1st November, 3rd November, and 5th November of 2025). Figure 1 shows the electricity P_t prices for each of the day used in the study.

The model parameters for the battery storage management were defined for a home-scale battery system with a storage capacity of $\overline{s} = 20.0$ kWh. The maximum potential of 4.6 kW which should be multiplied by one hour to provide

Table 1. Notation and used symbols

Indices	
t	slot (hour) of the day with fixed price, $t = 1, \ldots, T$
Data	
T	Last time period of the day
C	Cost of delivering back to the grid
P_t	Price of electricity per kWh in slot t
s_0	Stored energy at the beginning and at the end of the day
$\overline{s}$	Battery capacity
$\overline{b}$	Discharging capacity in one slot
$\overline{q}$	Charging capacity in one slot
Variables	
s_t	Amount stored in battery at the end of slot t in kWh
b_t	Quantity delivered back to grid in slot t in kWh
q_t	Quantity taken from grid in slot t in kWh

an upper bound $\overline{q} = 4.6$ kWh on the amount q_t that can be purchased in one hour from the grid. Similarly, the maximum discharging rate of 4.0 kW, which can be obtained from the battery, which translates to $\overline{b} = 4$ kWh as upper bound of what can be sold. Notice that one time slot has a duration of one hour, such that discharging at a rate of b_t during one time slot implies selling b_t kWh. The cost for selling back to the grid is set on $C = 0.02$ €/kWh and the storage at the end of the day is set to $s_0 = 15$.

3 Structure of the Model and Augmented Lagrangian

The presented model has a specific structure and can be reformulated as

$$\max_{y} b^\top y \quad \text{s.t. } Ay = c,\ 0 \leq y \leq \overline{y}.$$

Analyzing the corresponding matrix A, it consists of T rows defined by (3) and (2). The last row fixes the storage again at the end of the day $s_T = s_0$. Each row consists of 4 nonzero values $a_{ij} \in \{1, -1\}$, where the first and last equation have only 3 nonzero values (see Fig. 2). Basically, the matrix has dimension $T \times (3T - 1)$ and the number of non-zeros is $3 + 4(T - 2) + 3$ ($nz = 94$). Therefore, the matrix can be called sparse. Moreover, it is full row rank.

The structure of the resulting constraint matrix A shows significant sparsity, a feature that can be exploited by modern LP solvers and augmented Lagrangian algorithms [6]. Sparse representations are increasingly important for large-scale energy management problems, where thousands of time slots or nodes can be modeled efficiently [2].

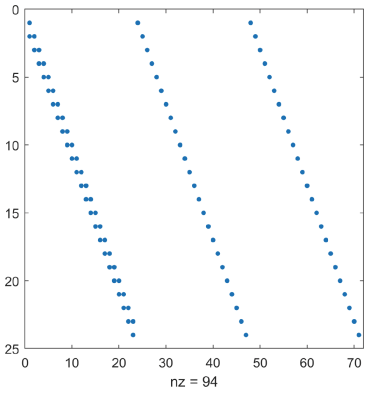


Fig. 2. Nonzero values in matrix A.

In [6], we developed a specific AL algorithm to deal with this type of problems in an efficient way. The advantage is that it has a first stage initialization to discover, if the given LP problem does not have a feasible point generating a so-called certificate of infeasibility of the dual. The second stage is an iterative scheme to generate an approximate basic solution in a finite number of steps [5,6]. The steps of the developed algorithm are depicted in Alg. 1.

Algorithm 1. Augmented Lagrangian

Require: A: matrix; c, $\overline{y}$, b: vectors; $\rho_0 > 0$: penalty parameter
$\omega > 1$: penalty updating factor; ϵ: tolerance; maxiter

$y_0 \in \operatorname{argmin}_{y \in [0,\overline{y}]} \frac{1}{2}(c - Ay)^\top (c - Ay)$
$d_0 = -(c - Ay_0)$
if $\|d_0\| \geq 10^{-6}$ **then** ▷ no feasible point exists
  return d_0
else ▷ y_0 is feasible, but not necessarily basic solution
  $k = 0$ ▷ Use y_0 as starting value for the AL optimization
  repeat
    $y_{k+1} \in \operatorname{argmin}_{y \in [0,\overline{y}]} \left(-b^\top y - d_k^\top (c - Ay) + \frac{\rho_k}{2}(c - Ay)^\top (c - Ay)\right)$
    $d_{k+1} = d_k - \rho_k(c - Ay_{k+1})$
    $\rho_{k+1} = \omega \rho_k$ ▷ update ρ
    $k = k + 1$
  until $\|y_k - y_{k-1}\| < \epsilon$ or $k >$maxiter
return y_k

4 Results and Discussion

The developed Algorithm 1 is an adaptation of the augmented Lagrangian method described in [6], in order to handle upper bounds (4)-(6) on the decision variables. The stopping criterion, based on the Euclidean norm, uses the

tolerance $\epsilon = \sqrt{\varepsilon_{\text{mach}}} = \sqrt{2^{-52}} \approx 1.49 \times 10^{-8}$ and the maximum number of iterations is set to maxiter = 120. According to [6], the initial value of the penalty parameter was set to $\rho = 1$, and is updated using $\omega = 3$.

The optimization in line 8 of Alg. 1 was implemented using the FMINCON solver from MATLAB®, which provides both flexibility in constraint handling and efficiency in convergence for sparse problems. The active set solver takes care of efficiently handling the bounds on the variables. We provide the gradient analytically to the solver. The termination tolerance was set to 10^{-12} for both y and the AL function value.

We provide Alg. 1 with a vector of zeros (zero dual infeasible point) as starting point for the solver in line 1. The results have been compared with the results of the LP solver LINPROG of MATLAB®. The reached objective function values are provided in Table 2.

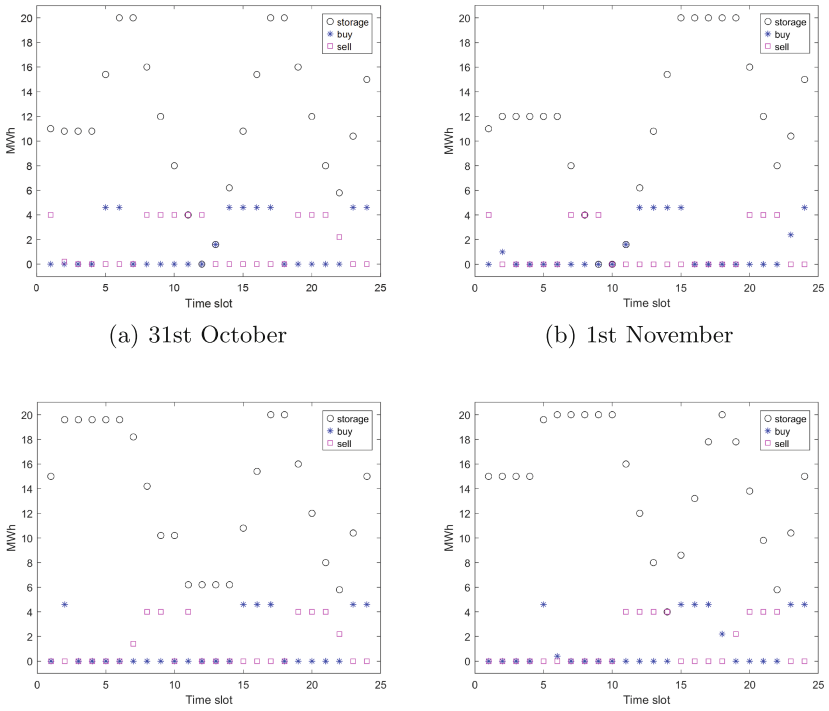
Table 2. Objective function value of LP and result of Alg. 1. The number of iterations (Iter) of Alg. 1 and the total number of function evaluations (Nr fev.) running the solver.

Day	Objective function LP	Objective function AL	Iter	Nr fev.
31st October	2.213558000000001	2.213557999999553	2	570
1st November	0.7690440000000003	0.7690439963065097	14	2467
3rd November	2.318600000000000	2.318599999872144	13	2092
5th November	1.765136000000000	1.765135999999051	3	649

The numerical results obtained with the energy values for each day are summarized in Table 2. The LP solution provides the optimum for the four price instances (days) and the resulting objective ranges from 0.7690 to 2.3186 €, corresponding to the net revenue achieved by the battery storage management of charging and discharging cycles. What can be observed is that the AL approach provides practically the same solution as the LP solver. It requires between 2 and 14 iterations, and fewer than 2500 function evaluations for all cases.

These results confirm the expected correlation between price volatility and achievable profit. On days with larger fluctuations in the hourly price P_t, the optimizer exploits opportunities for arbitrage - charging during low-price hours and discharging during peaks. For instance, the highest objective (2.3186 €) was obtained on 3rd November, when intraday price differences exceeded 0.10 €/kWh. The lowest return (0.7690 €) occurred on 1st November, when the price curve is more flat which provides less advantage in terms of providing revenue.

Figure 3 illustrates the energy movement and the storage profiles along the four days. The battery shows a cyclic feature, returning to its initial charge s_0 at the end of the day, to balance the daily storage. When the price rises, the system discharges energy to the grid (magenta curve). On the contrary, during off-peak hours, it draws power (blue curve) to replenish the storage.



(a) 31st October (b) 1st November

(c) 3rd November (d) 5th November

Fig. 3. Outcome of the model, in black the status of the battery, in blue the amount taken from the grid (bought) and in magenta the amount sold back to the grid. (Color figure online)

This behavior corresponds to the optimal scheduling reported in similar studies on LP-based and stochastic energy management frameworks [1,3]. The complexity of the model formulation just depends on the temporal resolution T. This means we can increase the density going for instance from an hour base to a 15 min base as is planned among Europe without increasing significantly the computational effort [2].

The numerical results show that the augmented Lagrangian method is able to find a feasible and near-optimal solution. Notice that the accuracy in this context can be in the order of euro cents. We found that the method requires a small number of iterations. Compared to classical simplex or interior-point solvers, the AL framework provides an interesting alternative. Notice that it is a first order method that does not depend on matrix inversions, see [5].

Notice that the described model and results focus on realistic price variation and bounds on operational activity to describe a real-world household battery behavior. The framework can be extended easily to include non-ideal characteristics such as round-trip efficiency, degradation cost, or temperature-dependent capacity as suggested in [1,4].

From a practical point of view, the described procedure shows that one can use sparse LP-based optimization for home or microgrid energy management systems. In such a practical environment, computation time and memory efficiency are relevant. Recent reviews [8] discuss that such formulations represent a promising step toward fully automated, cost-aware battery control in the context of dynamic electricity markets.

5 Conclusions

This study investigated the application of sparse linear programming and augmented Lagrangian methods to the battery storage management problem under variable electricity prices. The described model intends to optimize revenue in the context of a household-scale battery system. The idea is to exploit variability in real-world prices dynamics to maximize energy revenue by buying when the price is low and returning energy when the price is high. Due to the European trading system, these prices are known one day beforehand.

The numerical experiments show how a daily policy of buying and returning can be generated fast by a sparse LP formulation allowing future higher temporal discretization. We experimented with multiple daily price profiles. The model successfully adapted charging and discharging schedules to exploit market arbitrage, achieving higher profits on days with stronger price volatility.

One of our questions was whether an AL algorithm can handle effectively the upper-bound constraints in an all equality LP. The equalities describe the balance in the daily energy cycle, where we return to the initial storage at the end of each optimization horizon. We found that the AL approach being a first order method is efficient with respect to memory management and facilitates the use of iterative solvers. This shows that the AL approach can be a viable alternative to traditional simplex or interior-point methods in large-scale or embedded optimization contexts [5,6]. These findings are in line with recent findings in BESS optimization, where computational scalability and real-time applicability are becoming interesting challenges [2,8].

An interesting extension for future research of the current model to include the electricity consumption from the battery, efficiency losses, and temperature-dependent performance. The latter has shown to significantly affect long-term economic outcomes [1,4]. Another challenge is to include stochastic or predictive elements to represent uncertain renewable generation and price forecasts in the model to have a better robustness [3].

Overall, the results support the conclusion that sparse linear programming, combined with an augmented Lagrangian approach, provides a computationally efficient framework for battery storage optimization problem.

References

1. Bordin, C., Anuta, H.O., Crossland, A., Gutierrez, I.L., Dent, C.J., Vigo, D.: A linear programming approach for battery degradation optimization of off-grid systems. Renewable Energy **101**, 417–430 (2017). https://doi.org/10.1016/j.renene.2016.08.066
2. Chatzigeorgiou, N.G., Theocharides, S., Makrides, G., Georghiou, G.E.: A review on battery energy storage systems: applications, developments, and research trends of hybrid installations in the end-user sector. J. Energy Storage **86**, 111192 (2024). https://doi.org/10.1016/j.est.2024.111192
3. Engels, J., Claessens, B., Deconinck, G.: Combined stochastic optimization of frequency control and self-consumption with a battery. IEEE Trans. Smart Grid **10**(2), 1971–1981 (2017). https://doi.org/10.1109/TSG.2017.2785040
4. Gonzalez-Castellanos, A.J., Pozo, D., Bischi, A.: Non-ideal linear operation model for li-ion batteries. IEEE Trans. Power Syst. **35**(1), 672–682 (2019). https://doi.org/10.1109/TPWRS.2019.2930450
5. Guerrero-García, P., Hendrix, E.M.T.: Experiments with active-set LP algorithms allowing basis deficiency. Computers **12**(3), 1–23 (2023). https://doi.org/10.3390/computers12010003
6. Guerrero-García, P., Hendrix, E.M.T., Rocha, A.M.A.C.: On augmented lagrangian for LP feasibility. In: Gervasi, O., et al. (eds.) Computational Science and Its Applications – ICCSA 2025 Workshops, pp. 167–179. Springer Nature Switzerland, Cham (2026). https://doi.org/10.1007/978-3-031-97596-7_11
7. Sani Hassan, A., Cipcigan, L., Jenkins, N.: Optimal battery storage operation for PV systems with tariff incentives. Appl. Energy **203**, 422–441 (2017). https://doi.org/10.1016/j.apenergy.2017.06.043
8. Talebi, S., Aly, H.H.: A review of battery energy storage system optimization: current state-of- the-art and future trends. In: 2024 International Conference on Green Energy, Computing and Sustainable Technology (GECOST), pp. 116–120 (2024)

On Subproblem Solving in an Augmented Lagrangian Algorithm for LP with Upper Bounds

Eligius M. T. Hendrix[1(✉)], Pablo Guerrero-García[2], and Ana Maria A. C. Rocha[3]

[1] Computer Architecture, Universidad de Málaga, Spain and Wageningen University, 29071 Málaga, Spain
eligius@uma.es
[2] Applied Mathematics, Universidad de Málaga, 29071 Málaga, Spain
[3] Centro ALGORITMI/LASI, Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Portugal

Abstract. The augmented Lagrangian method has been well investigated for Linear Programming problems. We study LP problems where variables are nonnegative and have upper bounds. The resulting augmented Lagrangian subproblems require solving convex Quadratic Programming (QP) problems over a bound constrained area. We investigate the adaptation of three classical methods to solve the subproblems. We describe the algorithms and illustrate them with a small example.

Keywords: linear programming · augmented Lagrangian · bound constraints · quadratic programming

1 Introduction

Motivated by energy management problems [10], we are interested in Linear Programming (LP) problems with equality constraints and upper bounds on the variables. The upper bounds $\overline{x}$ on the variables in our application are caused by maximum storage and charging and discharging of batteries. In abstract mathematical terms this implies a problem of the form:

$$\max c^T x, \quad \text{s.t.} \quad Ax = b,\ x \in [0, \overline{x}] \subset \mathbb{R}^n, \tag{1}$$

where $c \subset \mathbb{R}^n$ the objective function gradient, b the right hand side, $\overline{x} \subset \mathbb{R}^n_+$ a vector of upper bounds and A a (sparse) $m \times n$ matrix.

Such a form includes upper bounds in what is called in several handbooks e.g. [1,2] the standard form LP; $\{\max c^T x | Ax = b, x \geq 0\}$. To avoid confusion,

This work has been supported by FCT Fundação para a Ciência e Tecnologia within the R&D Unit Project Scope UID/00319/2025 - Centro ALGORITMI (ALGORITMI/UM) https://doi.org/10.54499/UID/00319/2025.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 206–216, 2027.
https://doi.org/10.1007/978-3-032-30539-8_14

it should better be called the standard equality form. For this form, [3] and [4] investigated augmented Lagrangian (AL) algorithms to find a feasible point. Motivated by these studies, we constructed an AL algorithm to solve LP problem (1) with upper bounds on the variables and applied it in [10] comparing the results with the simplex method. In the application, we employed a generic constrained NLP algorithm to solve subproblems. However, specifically we are dealing with a bounded convex quadratic program. Our research question in this paper deals with an investigation of traditional methods that can solve these type of problems.

This paper is organized as follows. Section 2 introduces the setting of the augmented Lagrangian algorithm for the upper bounded LP. Sections 3, 4 and 5 discuss each a basic algorithm that can be adapted to solve the subproblems. Section 6 summarizes our findings.

2 Augmented Lagrange Embedding

We consider the upper bound extended equality LP model (1). By adding the linear constraints with Lagrange multipliers $\lambda \in \mathbb{R}^m$ to the objective, we obtain the equivalent Lagrangian problem

$$\begin{array}{l} \max_{0\le x\le \overline{x}}[c^T x, \quad \text{s.t.} \quad Ax = b] \\ \min_\lambda \max_{0\le x\le \overline{x}}[\mathcal{L}(x,\lambda) := c^T x + \lambda^T(Ax-b)] \\ \max_{0\le x\le \overline{x}}[c^T x + \min_\lambda\{\lambda^T(Ax-b)\}]. \end{array} \tag{2}$$

The analysis illustrates that if no feasible point x with $Ax = b$ exists, the Lagrangian is unbounded. The augmented Lagrangian approach we focus on, is to add a quadratic penalty to handle the equality constraints and iteratively solve the convex quadratic programming subproblem

$$\max_{0\le x\le \overline{x}}[c^T x + \lambda^T(Ax-b) - \frac{\rho}{2}(Ax-b)^T(Ax-b)], \tag{3}$$

where ρ is a penalty parameter which is updated during iterations. This problem is equivalent to a convex Quadratic programming (QP) problem, which we will write as a minimization problem

$$x^* \in \text{argmin}_{0\le x\le \overline{x}}[f(x) := r^T x + \frac{1}{2}x^T G x], \tag{4}$$

taking $G = \rho A^T A$ and $r = -A^T(\lambda + \rho b) - c$. Notice that the matrix $A^T A$ is positive semi-definite.

Algorithm 1 is the AL setting we investigate for optimizing (1). It first generates a feasible point x_0 by minimizing the square of the Euclidean norm of the slack. This is a so-called bounded least squares problem (LS). Notice that the LP has either no feasible point, one minimizer, or infinitely many; this corresponds to the argmin set. If no solution exists, then the minimum is not zero and the algorithm ends. Otherwise, the algorithm proceeds with an iterative process with x_0 and an estimate of the Lagrange multipliers λ_0 as starting point.

Algorithm 1. Augmented Lagrangian $(A, b, c, \overline{x}, \omega, \rho_0, \epsilon)$

Require: A: matrix
b, $\overline{x}$, c: vectors
$\rho_0 > 0$: penalty parameter, $\omega > 1$: update parameter, ϵ: accuracy
1: Find $x_0 \in \text{argmin}_{x \in [0,\overline{x}]} \frac{1}{2}(Ax-b)^T(Ax-b)$ ▷ bounded least squares
2: **if** $\|Ax_0 - b\| \geq 10^{-6}$ **then** ▷ no feasible point exists
3: **return** "No feasible point exists"
4: **else**
5: Initialize $k = 0, \rho_0 > 0$, $\lambda_0 = 0$ ▷ x_0 as starting value for the AL optimization
6: **repeat**
7: $x_{k+1} \in \text{argmax}_{x \in [0,\overline{x}]} \left(c^T x + \lambda_k^T (Ax-b) - \frac{\rho_k}{2}(Ax-b)^T(Ax-b)\right)$
8: $\lambda_{k+1} = \lambda_k - \rho_k(Ax_{k+1} - b)$
9: $\rho_{k+1} = \omega\rho_k$ ▷ update ρ
10: $k = k + 1$
11: **until** $\|x_k - x_{k-1}\| < \epsilon$
12: **return** x_k

In our preliminary experiments for the energy management problem using the AL algorithm, we solved subproblems using the FMINCON solver from MATLAB®, specifically the active set solver. However, such an approach is a generic algorithm for solving general NLP problems, whereas problem (4) has specific characteristics. Therefore, this paper investigates several algorithms from literature that are more specific and that can be adapted to solve (4), where positive semi-definite matrix G and vector r are given.

Exploiting the bound constraints, notice that the so-called Karush-Kuhn-Tucker (KKT) necessary conditions [6] for optimality in order to certify the necessary conditions for a found point to be an optimal solution of (4) are relatively easy. For a minimization problem, the negative of the objective function gradient should be a nonnegative linear combination of the gradients of the active constraints, e.g. [5]. For a bound constrained problem given by $x_j \in [0, \overline{x}_j]$, active (binding) constraints in an optimizer point x^* are characterized by $x_j^* = 0$ or $x_j^* = \overline{x}_j$, see also [7]. One can derive the following result.

Proposition 1. *Let x^* be a minimum point of (4) and define its objective function gradient as $g = Gx^* + r$. Then for $x_j^* = 0$, we have $g_j \geq 0$, for $x_j^* = \overline{x}_j$ we have $g_j \leq 0$ and for $0 < x_j^* < \overline{x}_j$ holds $g_j = 0$.*

We elaborate three specific algorithms from literature and illustrate them with a small instance. We first develop an algorithm to generate a solution of the bounded least squares (LS) problem in Line 1 of Alg. 1 to be presented in Sect. 3 called a projected gradient descent algorithm. We can adapt this algorithm to solve (4). In Sect. 4, this is followed by the description of a generic classical algorithm based on gradient projection and we adapt that to the box constraints. Finally, in Sect. 5 we study a specific gradient projection algorithm to solve a bound constrained QP (BCQP) problem. We first introduce a small example instance for which we sketch the steps of the algorithms.

Example 1. Consider an LP problem defined by

$$\max\{2x_1 + x_2\} \quad \text{s.t.} \quad 3x_1 + 2x_2 = 4,\ 0 \le x_j \le 1,\ j = 1, 2. \tag{5}$$

The optimal point is $x^* = (1, 0.5)^T$ with objective function value $c^T x^* = 2.5$. For the KKT conditions for this LP problem, it can be seen that the gradient c is a positive combination of the gradients of the active constraints $3x_1 + 2x_2 = 4$ and $x_1 \le 1$, specifically $(2, 1)^T = 0.5(3, 2)^T + 0.5(1, 0)^T$.

3 Projected Gradient Descent to Solve a Bounded LS

Although [9] speaks about projected gradient, what is projected is not the gradient, but a point following from a gradient step onto a bound constrained area. This is usually called projected gradient descent. We base our algorithm on [9] to find a solution of

$$\min_x \frac{1}{2}\|Ax - \beta\|^2, x_j \in [0, \overline{x}_j], j = 1, \ldots, n. \tag{6}$$

The algorithm in [9] is aiming at undetermined systems with $m \gg n$, where a solution $Ax = \beta$ does not exists, but should be approximated as good as possible. The problem to be solved in Line 1 of Algorithm 1 intends to find a feasible point of (1), where $n \gg m$. Notice that finding an optimal solution of (6) is equivalent to solving QP problem (4) where we take $G = A^T A$ and $r = -A^T \beta$.

The basic steps of the algorithm as depicted in Algorithm 2 is to use a step-size based on $L := \text{maxeig}(A^T A)$ in the direction $-\nabla f$. L is the maximum eigenvalue of $A^T A$. After taking the step, the result is projected on the box as feasible region. Projection here of a point z means that $P(z)$ takes component-wise:

$$P_j(z) = \max[\min(z_j, \overline{z}_j), 0].$$

Algorithm 2. Projected gradient descent $(A, \beta, \overline{x}, x_0)$

Require: A: matrix
β, $\overline{x}$: vectors, x_0: starting point
1: Compute L as $\text{maxeig}(A^T A)$
2: Initialize $k = 0$
3: **repeat**
4: $z = x_k - \frac{1}{L}\nabla f(x_k)$
5: $x_{k+1} = P(z)$
6: $k = k + 1$
7: **until** x_k is a stationary point or KKT point
8: **return** x_k

By taking $\beta = b$ and considering $G = A^T A$ and $r = -A^T b$ in (4), the gradient of $f(x)$ in (6) is given by

$$\nabla f(x) = A^T(Ax - b). \tag{7}$$

After using this gradient to find a feasible point in Line 1 of Algorithm 1, we can change the the quadratic objective in (4) and use it for iteratively solving the augmented Lagrangian steps in Line 7 of Algorithm 1, i.e.

$$\nabla f(x_k) = -A^T(\lambda_k - \rho_k b) - c + \rho_k A^T A x_k. \tag{8}$$

Example 2. We use the data of Example 1 setting $\beta = b$, compute $L = 13$ and set the initial value on $x_0 = (0,0)^T$. The corresponding least squares objective function value is $f(x_0) = \frac{4}{2} = 2$. The gradient in x_0 is $\nabla f(x_0) = -(12,8)^T$ providing the iterate $x_1 = \frac{1}{13}(12,8)$, which falls in the bound constraints and fulfills the equality constraint $Ax_1 = b$. This means that $\nabla f(x_1) = 0$, so it is a stationary point causing Algorithm 2 to stop.

We can now return to Algorithm 1 and use the outcome x_1 as feasible starting point $x_0 = \frac{1}{13}(12,8)$ in lines 6 to 11 setting initially $\rho_0 = 1$ and multiplier vector $\lambda_0 = 0$. For the AL optimization in Line 7, we now use the projected gradient descent Algorithm 2 using (8) to compute the gradient.

The gradient in x_0 is $\nabla f(x_0) = -(2,1)^T$, such that we want to increase both components of x. The gradient descent step leads to $z = x_0 - \frac{1}{13}\nabla f(x_0) = \frac{1}{13}(14,9)$, which is outside of the box. Projecting z provides point $(1, \frac{9}{13})^T \approx (1, 0.6923)^T$. This point is not a KKT point, because the second component of gradient is not zero. Notice that the KKT point the algorithm should converge to is $(1, 0.75)^T$. However, due to the fixed step size, it can be observed that the iterates approximate very slowly to this point. In fact, it takes the algorithm 92 iterations to reach a machine accuracy KKT point. Figure 1 shows 6 iterates of the algorithm.

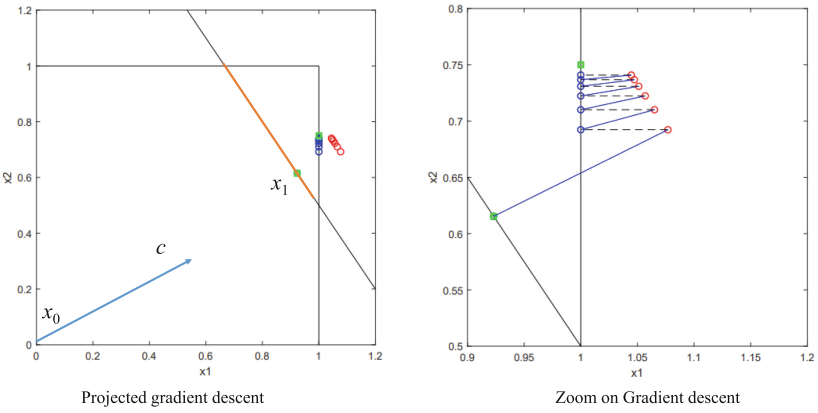


Fig. 1. At the left, the feasible set in orange with iterates of Algorithm 1 (Color figure online) as green squares. In red dots the gradient descent steps z, which are projected as blue points onto the feasible set. At the right a zoom and providing with (blue) lines the gradient descent steps and with black dashed lines the projection of 6 iterates.

As said, in one iteration a feasible point x_0 is reached. After that, the iterates zigzag between gradient descent and projection. This illustrates the slow convergence of the projected gradient descent algorithm.

The observation of slow convergence is also proven in theoretical sense; [9] shows that Alg. 2 converges with a speed proportional to $\frac{1}{k}$. Notice that such a convergence speed is less than linear.

4 A Gradient Projection on Linear Constraints

The original gradient projection concept was introduced by Rosen in 1960, see [11]. We investigate its adaptation to solve (4). It is based on projecting the gradient on the active constraints. In principle, this requires an inverse operation to project the gradient at each iteration where a new active constraint is found. Let M be the matrix of active constraints and g be the gradient in an iterate x_k, then

$$p = (I - M^T(MM^T)^{-1}M)g \tag{9}$$

is an increasing direction of the objective function in x_k. For minimizing the objective, the idea is to do a line-search in the direction of $-p$ taking care of the inactive constraints. Such a method appears unattractive when we are dealing with general linear constraints, as it requires a matrix inverse at each iteration. For problem (4), we are dealing with simple lower and upper bounds. That means, we can adapt the projection of the gradient in a point x taking component-wise:

$$p_j = \begin{cases} 0 & \text{if } x_j = 0 \text{ and } g_j > 0, \\ 0 & \text{if } x_j = \overline{x}_j \text{ and } g_j < 0, \\ g_j & \text{else.} \end{cases}$$

Stepping into the direction $-p$ along the line $x_k - \alpha p$, we have the directional derivative $dp = -r^T p - x_k^T G p$ for the convex quadratic objective. It may be clear that if $dp \geq 0$ apparently x_k cannot be improved in direction $-p$ and we found the best point. We adapt the generic approach to a quadratic objective which has the line-search as minimizing a quadratic one-dimensional function. This means we compute the directional second derivative $h = p^T G p$ and determine the corresponding step $\alpha^* = -\frac{dp}{h}$, taking care of $h > 0$. The maximum step size in Line 8 of Algorithm 3 is determined by the inactive constraints and calculated as

$$\overline{\alpha} := \min \left\{ \frac{x_j}{p_j} : p_j x_j > 0; \frac{x_j - \overline{x}_j}{p_j} : p_j(x_j - \overline{x}_j) > 0 \right\}. \tag{10}$$

Example 3. We take from Example 2 the reached feasible point $x_0 = \frac{1}{13}(12, 8)$ to start the optimization of Algorithm 1 Line 7, now using Algorithm 3. This means, we have as input $G = \rho_0 A^T A = \begin{pmatrix} 9 & 6 \\ 6 & 4 \end{pmatrix}$ and $r = \begin{pmatrix} -14 \\ -9 \end{pmatrix}$, such that $g = -(2, 1)^T$. As no box constraint is active, the projected gradient is $p = g$. In

Algorithm 3. Rosen gradient projection $(G, r, \overline{x}, x_0)$

Require: G: positive semi-definite matrix
$r, \overline{x}$: vectors, x_0: feasible starting point
Initialize $k = 0$
repeat
  $g = Gx_k + r$
  Determine projected gradient p and maximum step size $\overline{\alpha}$ according to (10)
  Compute directional derivative $dp = -(r^T p + x_k^T Gp)$
  Compute second derivative $h = p^T Gp$ and step-size $\alpha^* = \frac{-dp}{h}$
  if $dp < 0$ **then**
    $x_{k+1} = x_k - \min\{\alpha^*, \overline{\alpha}\}p$
  else
    $x_{k+1} = x_k$
  $k = k + 1$
until x_k is a stationary point or KKT point
return x_k

order not to get out of the box in direction $-g$, the maximum step is $\overline{\alpha} = \frac{1}{26}$, whereas $\alpha^* = \frac{5}{64}$. The corresponding next iterate $x_1 = (1, \frac{17}{26})^T \approx (1, 0.6538)^T$ is not a KKT point. Notice the numerical difference with x_1 in the projected gradient descent approach (Algorithm 2), which projects after taking a step. As the first component of x_1 is on the upper bound and both derivatives are negative, the projected gradient only contains the second component as nonzero, $p = -(0, 0.3846)^T$. Line-search to determine the step-size α^* provides the KKT point $x_2 = (1, 0.75)^T$. This is the optimum of the AL function for $\rho_0 = 1$ and $\lambda_0 = 0$ in Algorithm 1. The iterates are depicted in Fig. 2. In green are the main iterates of the AL algorithm.

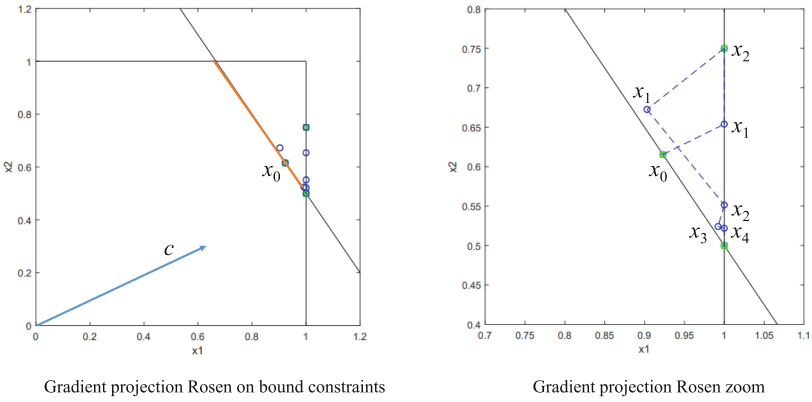


Fig. 2. Feasible set in orange with AL iterates of Algorithm 1 (Color figure online) in green squares. At the left, iterates of Algorithm 3 in blue. At the right, a zoom and the steps depicted by lines.

Following Algorithm 1, we should now update Lagrange multiplier $\lambda_1 = \lambda_0 - \rho_0(Ax_1 - b) = -0.5$ and increase $\rho_1 = 2$. We call Algorithm 3 again for Line 7 of Algorithm 1 with input $G = \rho_1 A^T A = \begin{pmatrix} 18 & 12 \\ 12 & 8 \end{pmatrix}$ and $r = \begin{pmatrix} -24.5 \\ -16 \end{pmatrix}$. So, $g = \begin{pmatrix} 2.5 \\ 2 \end{pmatrix}$. Algorithm 3 now requires 5 steps to reach the KKT point $x^* = (1, 0.5)^T$ which coincides with the optimal point of the LP problem.

5 A Gradient Projection Method to Solve a BCQP

We focus on a gradient projection method for the bound constrained quadratic programming (BCQP) problem outlined in [8]. The initial ideas can be found in [7]. It has been developed to find specifically a minimum point of $\frac{1}{2}x^T Gx + r^T x$ over bound constraints. Substituting $G = \rho A^T A$ and $r = -A^T(\lambda + \rho b) - c$, it may be used to solve Line 7 in Algorithm 1. In contrast to the method introduced by Rosen (see Sect. 4), the search direction is bent along the active box constraints to compute a so-called Cauchy point such that it stays in the box constrained area. Only after a minimum point is found along the search trajectory, we compute a new gradient. To describe the trajectory, [7] introduced the concept of breakpoints. Its computation can be derived from the description in [8].

The algorithm is depicted in Algorithm 4. In addition to the usual iterates x_k we have an inner loop with iterates z_i along the breakpoints of the trajectory starting at x_k. The algorithm follows iterative gradient driven trajectories up to finding an optimal point on them. It identifies breakpoints t_j in the direction of the negative gradient where bound constraints become active. Given the gradient in $g = Gx + r$ in iterate x, we find for each component $j = 1, \ldots, n$ the breakpoint

$$t_j = \begin{cases} \frac{x_j}{g_j} & \text{if } g_j > 0, \\ \frac{x_j - \bar{x}_j}{g_j} & \text{if } g_j < 0, \\ \infty & \text{else.} \end{cases} \tag{11}$$

We order the set of breakpoints removing duplicates starting (potentially adding) at 0 and ending at ∞, to generate the list $\tau_0 = 0, \tau_1, \tau_2, \ldots, \infty$. The ordered τ_i define a piecewise linear trajectory.

Each line-piece $[0, \tau_1], [\tau_1, \tau_2], \ldots, [\tau_i, \tau_{i+1}]$ corresponds to a (negative) projected gradient vector q with elements

$$q_j = \begin{cases} -g_j & \text{if } t_j > \tau_i, \\ 0 & \text{else.} \end{cases} \tag{12}$$

Starting at $z_0 = x_k$, the algorithm walks along the intervals. In the interval from point z_i to next point $z_{i+1} = z_i + (\tau_{i+1} - \tau_i)q$, we look for the minimum of the quadratic function. Define the directional derivative $dq = r^T q + z_i^T Gq$. It may be clear that if $dq \geq 0$, apparently z_i cannot be improved in direction q and we found the best point. If we can still improve, we compute the directional

Algorithm 4. Trajectory gradient projection for BCQP $(G, r, \overline{x}, x_0)$

Require: G: positive semi-definite matrix
$r, \overline{x}$: vectors, x_0: starting point
1: Initialize $k = 0$
2: **repeat**
3: $g = Gx_k + r$
4: Determine breakpoints $t_1, \ldots, t_n$ according to (11)
5: Obtain ordered line-piece intervals $[0, \tau_1], [\tau_1, \tau_2], \ldots$
6: Set $i = 0$, next_interval = true and $z_0 = x_k$ ▷ start the trajectory
7: **while** next_interval **do**
8: Compute direction q given τ_i according to (12)
9: Compute directional derivative $dq = r^T q + z_i^T Gq$
10: **if** $dq \geq 0$ **then** ▷ minimum for τ_i
11: $x_{k+1} = z_i$ and next_interval = false
12: **else**
13: Compute second derivative $h = q^T Gq$ and step $\alpha^* = \frac{-dq}{h}$
14: **if** $\tau_i + \alpha^* \leq \tau_{i+1}$ **then** ▷ interior minimum over line-piece
15: $x_{k+1} = z_i + \alpha^* q$ and next_interval = false
16: **else** ▷ look into next interval
17: $z_{i+1} = z_i + (\tau_{i+1} - \tau_i)q$
18: $i = i + 1$
19: $k = k + 1$
20: **until** x_k is a stationary point or KKT point
21: **return** x_k

second derivative $h = q^T Gq$ and determine the corresponding step $\alpha^* = -\frac{dq}{h}$. If $\alpha^* \leq \tau_{i+1}$ we take as minimum point $z_i + \alpha^* q$ and stop. Otherwise we go to the next interval $[\tau_{i+1}, \tau_{i+2}]$ to find a better point. The complete algorithm is sketched in Algorithm 4. Notice that [7] proved theoretically that under mild conditions, the algorithm converges in a finite number of steps.

We elaborate the algorithm for the example instance where it is used in the context of solving problem (3) in an iterative way.

Example 4. Like in Example 3, we take the outcome of Algorithm 2 $x_0 = \frac{1}{13}(12, 8)$, as starting point and for Algorithm 1 we set $\rho = 1$. This means the initial input of Algorithm 4 is $G = \rho_0 A^T A = \begin{pmatrix} 9 & 6 \\ 6 & 4 \end{pmatrix}$ and $r = \begin{pmatrix} -14 \\ -9 \end{pmatrix}$. Consequently, $g = \begin{pmatrix} -2 \\ -1 \end{pmatrix}$, such that we want to increase both components of x. In order not to get out of the box in direction $-g$ we get breakpoints $t_1 = \frac{1}{26}$ and $t_2 = \frac{5}{13}$. The minimum is found along the piecewise linear trajectory in between those breakpoints providing iterate $x_1 = (1, 0.75)$. Point $x^* = x_1$ is a KKT point according to Prop. 1, as $g = Gx^* + r = (-0.5, 0)^T$, i.e. $g_1 < 0$ for $x_1^* = \overline{x}_1$ and $g_2 = 0$ for $0 < x_2^* < \overline{x}_2$.

The outcome $x_1 = (1, 0.75)^T$ of the gradient projection approach for $\rho_0 = 1$ provides us with a minimum point of Line 7in Algorithm1. Following the

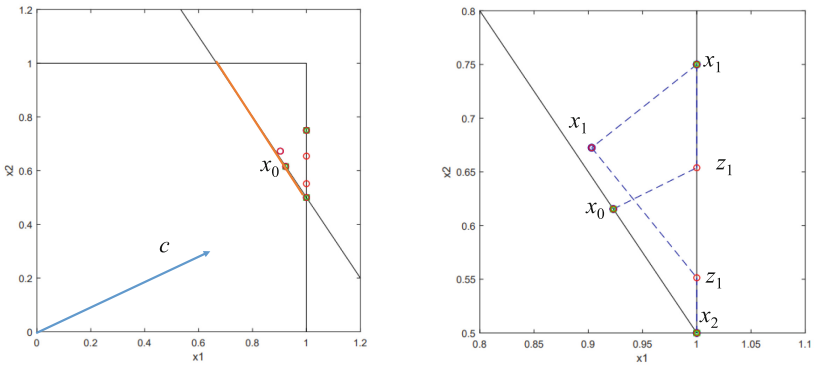


Fig. 3. Feasible set in orange with AL iterates of Algorithm 1 (Color figure online) in green squares. At the left, iterates of Algorithm 4 in blue. At the right, intermediate points z_j of Algorithm 4 in red dots.

iterations of Algorithm 1, we compute the update of the Lagrange multiplier $\lambda_1 = \lambda_0 - \rho(Ax_1 - b) = -0.5$ and increase $\rho_1 = 2$. This provides input G and r like in Example 3.

To decrease both components of x gives breakpoints $t_1 = \frac{2}{5}$ and $t_2 = \frac{3}{8}$. This provides an interior optimum over the first line-piece of $x = (0.9031, 0.6725)^T$ close to the feasible set of the LP with gradient $g = (-0.1739, 0.2173)^T$. Direction $q = -g$ goes in parallel along the equality constraint and leads to point $z_1 = (1, 0.5514)^T$. From there it bends the directory towards $z_2 = x_2 = (1, 0.5)^T$. This point corresponds to the optimal solution x^* of the original LP problem. The method of Rosen requires a few more steps in the end to reach the optimum. The difference is that it computes the gradient in each intermediate point. The steps of the AL algorithm and the intermediate trajectories are depicted in Fig. 3, where intermediate points z_i are given in red.

6 Conclusions

This paper focuses on an augmented Lagrangian approach to solve a linear programming equality problem with upper bounds. Such LP problems are relevant in energy management. It describes three algorithms that can be used to solve the bound constrained quadratic programming (BCQP) subproblem in the context of the augmented Lagrangian algorithm.

First we present a least squares projected gradient descent algorithm adapted to solve the subproblem. The second algorithm is the classical algorithm of Rosen adapted to the box constraints using an explicit line search. Finally, we study a specific gradient projection algorithm to solve a BCQP problem. A very small instance with only one equality illustrates some characteristics of the methods.

We found that the projected gradient descent method converges extremely slow due to the fixed step size. The classical gradient projection approach

of Rosen, solves the problem in a few iterations thanks to exploiting a one-dimensional quadratic line-search to determine the step-size. The specific BCQP approach also solves the small instance fast, but its scheme is more complicated due to following a piecewise linear trajectory. It requires even less steps for the small example than the gradient projection of Rosen. The computational burden of the last two approaches is relatively low.

In future research, we will study their behavior on practical instances.

References

1. Bertsimas, D., Tsitsiklis, J.N.: Introduction to Linear Optimization. Athena Scientific, Belmon, Massachusetts (1997)
2. Claassen, G.D.H., Hendriks, T.H.B., Hendrix, E.M.T.: Decision Science. Wageningen Academic Publishers, Wageningen (2007)
3. Galabova, I., Hall, J.: The 'idiot' crash quadratic penalty algorithm for linear programming and its application to linearizations of quadratic assignment problems. Optimization Meth. Softw. **35**(3), 488–501 (2020). https://doi.org/10.1080/10556788.2019.1604702
4. Guerrero-García, P., Hendrix, E.M.T., Rocha, A.M.A.C.: On augmented Lagrangian for LP feasibility. In: Gervasi, O., et al. (eds.) Computational Science and Its Applications – ICCSA 2025 Workshops, pp. 167–179. Springer Nature Switzerland, Cham (2026). https://doi.org/10.1007/978-3-031-97596-7_11
5. Hendrix, E.M.T., Tóth, G., B.: Introduction to Nonlinear and Global Optimization, Springer (2010)
6. Kuhn, H.W., Tucker, A.W.: Nonlinear programming. In: Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950, pp. 481–492. Univ. California Press, Berkeley-Los Angeles, Calif. (1951)
7. Moré, J.J., Toraldo, G.: Algorithms for bound constrained quadratic programming problems. Numerische Mathematik **55**(4), 377–400 (1989). https://doi.org/10.1007/BF01396045
8. Nocedal, J., Wright, S.J.: Numerical Optimization, 2e edn. Springer, New York, NY, USA (2006)
9. Polyak, R. A.: Projected gradient method for non-negative least square. Contemp. Math. **636**, 167–179 (2015). https://doi.org/10.1090/conm/636/12735
10. Rocha, A.M.A.C., Guerrero-García, P., Hendrix, E.M.T.: On battery storage management by sparse linear programming. In: Gervasi, O., et al. (eds.) Computational Science and Its Applications – ICCSA 2026 Workshop. Springer Nature Switzerland, Cham (2026). https://doi.org/10.1007/978-3-032-30539-8_13
11. Rosen, J.B.: The gradient projection method for nonlinear programming. part I. linear constraints. J. Soc. Indust. Appl. Math. **8**(1), 181–217 (1960). https://doi.org/10.1007/978-3-032-30539-8_13

Resilient Mobility Under Multiple Shocks: The SCUDO Project Experience

Francesco Bruzzone(✉), Matilde Gasparetto, and Silvio Nocera

Università Iuav di Venezia, Santa Croce 191, 30125 Venice, Italy
fbruzzone@iuav.it

Abstract. The transport sector is increasingly exposed to disruptive events such as pandemics, natural hazards, cyber-attacks, and energy crises, which undermine both its reliability and its contribution to social equity. Traditional approaches have largely focused on protecting physical infrastructure, while more recent understandings of resilience emphasize the interdependence of technical performance, governance capacity, and users' socio-economic conditions. This paper develops an integrated framework of Key Performance Indicators (KPIs) that connects system-level resilience with socio-economic vulnerability to assess how different population groups anticipate, absorb, and recover from transport disruptions. The methodology combines a structured literature review with indicator design across six dimensions: robustness of infrastructure, operational continuity, governance and institutions, socio-economic accessibility, environmental sustainability, and systemic coherence. The framework is stress-tested on three types of disruption: pandemic, flood-related natural disasters, and energy price shocks-drawing on EU-level data and previous empirical studies to link performance metrics with profiles of vulnerable users. Results highlight that low-income, carless or car-dependent commuters, digitally excluded workers, and residents of isolated rural areas experience disproportionate exposure to service interruptions and energy price volatility, and benefit least from adaptive mechanisms such as telework, multimodal alternatives, and digital decision support. The paper argues that embedding equity-sensitive, resilience-oriented KPIs in transport planning and evaluation can support more robust, inclusive, and climate-resilient mobility policies across Europe.

Keywords: Transport disruption · Transport vulnerability · Key Performance Indicators · Natural disasters

1 Introduction

In recent years, transport systems have faced mounting disruptions from pandemics, extreme weather events, cyber incidents, and energy-related shocks, challenging both operational reliability and the sector's contribution to social equity [1]. These shocks

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 217–225, 2027.
https://doi.org/10.1007/978-3-032-30539-8_15

increasingly interact with long-term pressures such as climate change, demographic ageing and digitalization, creating complex risk profiles that cannot be adequately understood through purely engineering-based approaches [2]. At the same time, decarbonization pathways and the electrification of transport infrastructures introduce new dependencies on energy and digital systems, which may themselves become sources of vulnerability under disruption [1]. Traditional resilience strategies have typically prioritized the protection and rapid restoration of physical infrastructure, emphasising standards, design codes and asset management practices aimed at reducing direct damage and shortening repair times. While these perspectives remain essential, they often treat users as homogeneous and pay limited attention to how different social groups experience and respond to disruptions in terms of accessibility, affordability and travel opportunities. In addition, most existing indicator frameworks assess single hazards or focus on network-level performance, with limited capacity to capture cascading effects across systems and distributional consequences for vulnerable groups. Recent approaches argue that resilience should instead be conceptualized as a socio-technical property, emerging from interactions between infrastructure robustness, operational strategies, institutional capacities, and the resources and constraints that shape user behavior [2]. In this perspective, there is a growing need for integrated indicator sets that link these interacting dimensions to concrete disruption archetypes and vulnerability profiles, in a way that is operational for planning and evaluation. Building on early evidence and methodological work developed within the SCUDO (Safe and Comprehensive Urban Development and Transport management for Disruption-resilient Operations) project, this paper proposes an integrated KPI (Key Performance Indicators) framework that explicitly connects disruption types, system responses, and vulnerability profiles across different territorial contexts. The framework responds to calls in the literature for metrics that can support scenario-based assessment of resilience measures while avoiding "average-only" evaluations that obscure distributional effects. It does so by combining quantitative indicators of network performance and recovery with qualitative indicators of governance, institutional coordination and user experience, structured in a way that allows decision-makers to trace how shocks propagate through systems of infrastructure, services and social practices.

2 Conceptual Framework

A growing body of literature conceptualizes resilience as a combination of outcomes (such as continuity of service and time-to-recovery) and capabilities (such as preparedness, coordination, adaptability, and learning) [3, 4]. Outcome-oriented indicators typically measure how much service is maintained during a disruption, how quickly performance is restored, and how severe the residual impacts are for users and operators [5]. Capability-oriented indicators, by contrast, focus on the institutional and organizational arrangements that enable systems to anticipate shocks, absorb disturbances, reorganize operations and incorporate lessons into future planning cycles [4]. Bringing these two perspectives together allows transport authorities to evaluate not only whether systems "bounce back", but also whether they can adapt and transform in ways that reduce future exposure and vulnerability [6].

Evidence syntheses show that disruptions can generate cascading impacts on mobility and accessibility, with uneven consequences across socio-economic groups and places.

For instance, damage to critical links or power supply can propagate through public transport and logistics networks, affecting not only travel times but also access to employment, health care and essential goods, especially in territories with limited redundancy or weak multimodal integration [1, 2]. The literature further demonstrates that the same disruption can have markedly different implications in dense urban cores, peri-urban belts and sparsely populated rural regions, depending on baseline service provision, availability of alternatives and the spatial distribution of opportunities [2].

Low-income households, car-less users, older adults, people with disabilities, women and caregivers, migrant workers, and rural residents frequently face higher exposure to disruptions, fewer feasible alternatives, and slower recovery trajectories [1, 5]. These groups are often constrained by limited financial resources, job inflexibility, caring responsibilities, physical or sensory impairments, language barriers and restricted access to digital information, all of which reduce their capacity to reroute, reschedule or switch modes when regular services are disrupted [2, 4]. Non-teleworkable workers, for example, remain dependent on commuting during pandemics or fuel price spikes, while many rural residents have no viable alternatives to a small number of bus routes or road corridors when these are affected by floods or snowstorms. Recognizing these differentiated vulnerabilities is therefore central to any attempt to design resilience metrics that go beyond system averages and account for distributional outcomes.

Recent contributions emphasize that bridging human and infrastructure resilience requires metrics that capture both network performance and the social capacity to anticipate, absorb, and adapt to shocks [2]. This implies integrating insights from socio-ecological resilience theory, which highlights feedbacks between human behavior, institutions and physical systems, with more traditional reliability and risk assessment approaches used in transport engineering [3]. From this vantage point, resilience is not an inherent property of assets or technologies, but the emergent result of how they are planned, governed and used by diverse actors over time (Fig. 1) [4]. The KPI framework proposed in this paper builds on this understanding, seeking to move beyond “average-only” solutions toward equity-sensitive resilience appraisal that explicitly links disruption types, performance indicators, governance capacities and user vulnerabilities within a coherent analytical structure.

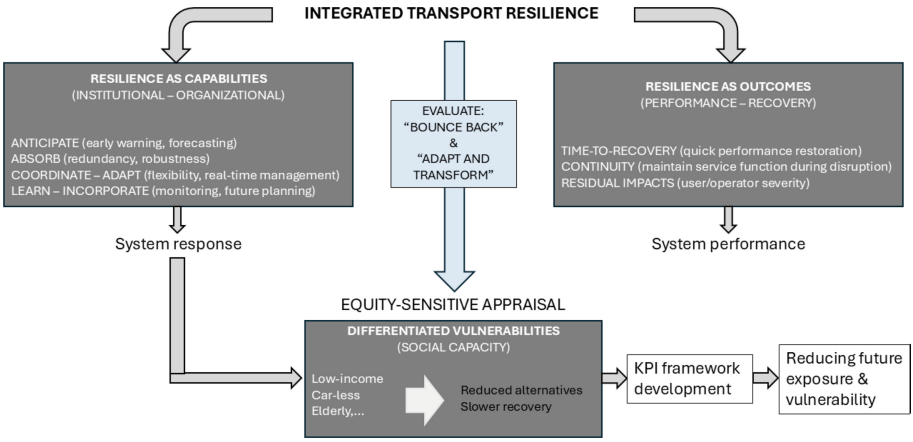


Fig. 1. Overview of the conceptual framework. Own elaboration.

3 Methodology

The methodology combines a structured review of peer-reviewed studies and EU reports—see Sect. 2—with the design of an operational indicator framework for resilience appraisal (Fig. 2).

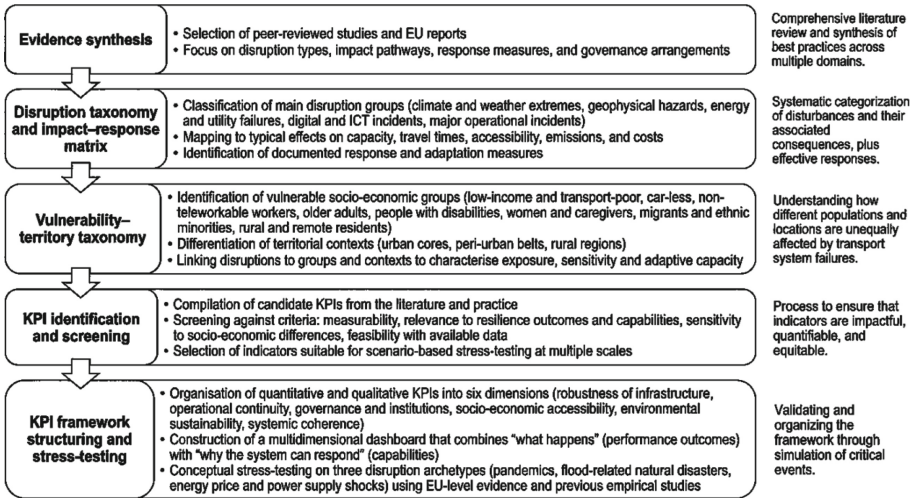


Fig. 2. Methodological framework (own elaboration)

The literature review is organized in three steps: first, it identifies key conceptual and empirical contributions on transport resilience, focusing on how disruptions affect both infrastructure performance and user experience. Second, it extracts information

on disruption types, impact pathways, and response measures, paying particular attention to the role of institutions, governance arrangements and recovery strategies. Third, it synthesizes this material into a set of attributes and indicators that can be used to characterize resilience outcomes and capabilities across different transport modes and territorial contexts.

The main groups of transport disruptions are classified, including climate and weather extremes (floods, storms, heatwaves, wildfires), geophysical hazards (earthquakes), energy and utility failures (power supply disruptions), digital and ICT incidents (cyber disruptions), and major operational incidents (accidents, infrastructure failures, bottlenecks). For each disruption type, the review traces impact pathways on mobility services and accessibility, along with response measures and governance arrangements documented in case studies, discussion papers and technical guidelines. This allows the construction of a disruption/impact/response matrix that links hazards to typical effects on capacity, travel times, accessibility, emissions and costs, as well as to common mitigation and adaptation strategies.

The material is then used to develop an integrated taxonomy that links disruptions to socio-economic groups and geographic contexts, enabling KPI selection and scenario design that capture both performance and distributional outcomes. The taxonomy distinguishes, for example, between dense urban corridors with multiple modes, peri-urban areas with car-dependent travel patterns, and rural regions with sparse services, and associates each configuration with specific profiles of vulnerability and adaptive capacity. On this basis, candidate KPIs are screened against criteria of measurability, relevance to resilience outcomes and capabilities, sensitivity to socio-economic differences, and feasibility of application with available or realistically collectable data.

The resulting framework organizes quantitative and qualitative indicators around six dimensions, building on Karagiannakis et al. [7]: robustness of infrastructure, operational continuity, governance and institutions, socio-economic accessibility, environmental sustainability, and systemic coherence. Quantitative KPIs capture aspects such as redundancy ratios, percentage of service delivered, delay and cancellation rates, energy recovery intensity and composite resilience indices. Meanwhile, qualitative KPIs characterize, among others, coordination structures, equity criteria, integration with decarbonization policies and the role of digitalization in supporting decision-making and communication. These performance indicators form the backbone of a dashboard that can be applied in stress-testing exercises, where different disruption scenarios are used to explore how changes in infrastructure, operations or governance affect both system-level performance and the conditions of vulnerable user groups across diverse territorial settings.

Beyond defining the indicator set, the methodology explicitly addresses the sensitivity and reliability of the proposed KPIs before full application to a specific local case. Sensitivity is supported, first, by grounding the disruption/impact/response matrix in empirical evidence on how different shocks affect capacity, travel times, accessibility, emissions and costs, and second, by selecting indicators that have demonstrated responsiveness under disruption conditions in previous studies, such as percentage of service delivered, delay and cancellation rates, and accessibility losses for specific user groups.

Reliability is strengthened through screening criteria that prioritise measurability, interpretability for decision-makers, and feasibility with EU-level and national data, while ensuring that indicators remain applicable across diverse network structures and territorial contexts. The KPI framework is thus conceived as a transferable template which reflects consolidated evidence on disruption mechanisms and vulnerability patterns in different transport systems and modes.

4 KPI Framework Design and Testing

The KPI framework (Table 1) assesses resilience through a dashboard that combines "what happens" (performance outcomes) with "why the system can respond" (capabilities) [8]. Robustness of infrastructure includes indicators on the structural resistance of assets, redundancy in critical links, and vulnerability of nodes and corridors to specific hazards such as flooding or earthquakes. Operational continuity comprises metrics on service availability, capacity utilization, speed of degradation under stress, and time-to-recovery of different modes and routes. Governance and institutional capacity is captured through indicators of preparedness (existence of contingency plans and early warning systems), coordination across agencies and operators, data and information-sharing practices, and stakeholder involvement in resilience planning. Socio-economic accessibility focuses on the ability of different population groups to reach essential activities under disruption conditions, through indicators of affordability, availability of alternatives, and adaptability of travel behavior. Environmental sustainability incorporates metrics related to emissions, energy efficiency, and potential trade-offs between short-term recovery measures and long-term decarbonization goals. Finally, systemic coherence examines alignment across infrastructure, operations, governance, and user needs, ensuring that measures designed to protect network performance do not inadvertently deepen inequalities or lock in unsustainable patterns. In combination with the disruption/impact/response matrix and the disruption-vulnerability taxonomy developed in Sect. 3, this framework therefore offers an ex-ante tested structure in which indicators are expected to react in a consistent and interpretable way under different disruption scenarios.

The framework is stress-tested on three disruption archetypes: pandemics, flood-related natural disasters, and energy price or supply shocks, drawing on EU-relevant evidence and previous empirical studies [1, 5, 11]. Pandemics alter demand patterns through mobility restrictions and telework adoption, while imposing capacity constraints and safety protocols on public transport, with unequal impacts on workers whose jobs cannot be performed remotely. Flood-related disasters damage infrastructure, disrupt operations, and reconfigure accessibility landscapes, especially in territories with limited redundancy and weak evacuation logistics. Energy price shocks and power supply disruptions affect operating costs, service frequency, and fare levels, with knock-on effects on affordability and mode choice, particularly for households with constrained budgets or high commuting dependence.

Across these scenarios, several socio-economic groups consistently exhibit higher levels of vulnerability. Low-income and transport-poor households are more exposed to affordability stress and possess fewer options to substitute disrupted services or

Table 1. The KPI framework

Dimension	Quantitative KPIs	Qualitative KPIs	Sources
Technical / Infrastructure	Redundancy ratio; MTTR; % capacity loss	Resilience standards/codes	[7, 9]
Operations / Service performance	% service delivered; delay/cancellation rate; rerouting speed	Operational flexibility; passenger information procedures	[3, 5]
Governance / Institutional capacity	Plan update frequency; resilience budget share; joint exercises	Coordination structures; stakeholder inclusion	[4, 10]
Users & Equity	Access to essential services; affordability stress; coverage of vulnerable groups	Equity criteria; perceived reliability; inclusive communication	[2, 5]
Environment / Sustainability	Net emissions avoided; recovery energy intensity	Integration with decarbonisation policies	[1, 6]
Systemic / Cross-cutting	Composite resilience index; stress-test score	SDG alignment; electrification/digitalisation role	[8, 11]

absorb fares and fuel increases. Car-less households depend strongly on public transport and often lack accessible alternative modes when services are reduced or rerouted. Elderly people and people with disabilities face heightened sensitivity to disruption and additional accessibility barriers during emergency responses or recovery works. Women and caregivers are constrained by complex trip-chaining and care responsibilities, which make even minor disruptions disproportionately burdensome. Migrants and ethnic minorities may encounter language and information barriers, as well as rigid job and housing constraints that reduce adaptive capacity. Non-teleworkable workers are highly dependent on commuting even under mobility restrictions, limiting their ability to benefit from digital-based mitigation strategies. Rural and remote residents suffer from lower baseline service availability and longer access distances, leaving them more exposed to both service interruptions and energy price volatility.

By mapping disruption types to vulnerable groups and territorial configurations, the KPI framework supports scenario-based assessments that reveal both technical and equity implications of resilience strategies. Multidimensional KPIs enable ex-ante comparison of measures across infrastructure reinforcement, operational reconfiguration, governance reforms, and user-oriented interventions, thereby reducing reliance on single-metric or average performance evaluations [6]. For example, indicators of socio-economic accessibility and affordability can be combined with service continuity metrics to identify where maintaining minimum service levels for specific groups may be more critical than optimizing overall network efficiency under shock conditions.

Similarly, governance indicators related to communication channels, participatory planning, and data inclusiveness help assess whether digital tools and decision-support systems are accessible to digitally excluded users. The framework thus connects resilience enhancement with the design of targeted support mechanisms, such as protected service corridors for key workers, subsidized fares for low-income commuters, or enhanced demand-responsive transport in underserved rural areas [1, 5].

5 Discussion and Conclusion

The analysis confirms that resilience-oriented KPIs must systematically account for how different population groups anticipate, absorb, and recover from disruptions, paying particular attention to those with limited financial resources, constrained modal choices, or restricted digital access. The results highlight that low-income user, car-less or car-dependent commuters, digitally excluded workers, and residents of isolated rural areas experience disproportionate exposure to service interruptions and energy price volatility, and often benefit least from adaptive mechanisms such as telework, multimodal alternatives, or real-time digital information. Equity-sensitive KPI dashboards can guide policymakers in interrogating trade-offs between short-term restoration of capacity and longer-term goals of climate resilience and social inclusion [4, 6]. Embedding such dashboards into planning and evaluation routines may also facilitate cross-sectoral coordination, aligning transport resilience strategies with social policy, energy transitions, and regional development objectives [1, 10].

In conclusion, this paper presented an integrated KPI framework that connects system-level resilience with socio-economic vulnerability, structured around six dimensions spanning infrastructure, operations, governance, accessibility, environment, and systemic coherence. Drawing on evidence that disruptions generate cascading and unequal impacts, the framework foregrounds the experience of vulnerable users and enables equity-sensitive scenario design and appraisal. The approach is rooted in methodological advances initiated within the SCUDO project but is conceived to be transferable across European contexts and adaptable to different types of disruption. By embedding equity-sensitive, resilience-oriented KPIs into transport planning and evaluation, policymakers can support more robust, inclusive, and climate-resilient mobility systems under conditions of growing uncertainty and multiple shocks. Future research should refine indicator sets with empirical data, test their performance in real-world decision processes, and further explore how institutional arrangements shape both resilience outcomes and the distribution of benefits across user groups.

Acknowledgments. This research was funded by Horizon Europe SCUDO (*Safe and Comprehensive Urban Development and Transport management for Disruption-resilient Operations*) Project.

Disclosure of Interests The authors declare no competing interest.

References

1. Glock, J.P., Hartl, R., Krieg, M., Becker, U.J.: No climate-resilient society without a resilient transport system. In: The Palgrave Handbook of Climate Resilient Societies, vol. 1–2, pp. 1147–1175. Springer (2022)
2. Hayes, S., Desha, C., Burke, M., Gibbs, M., Chester, M.: Leveraging socio-ecological resilience theory to build climate resilience in transport infrastructure. Transp. Rev. **39**(5), 677–699 (2019)
3. Chan, R., Schofer, J.L.: Measuring transportation system resilience: response of rail transit to weather disruptions. Nat. Hazards Rev. **17**(1), 05015004 (2016)
4. Renne, J., Wolshon, B., Murray-Tuite, P., Pande, A.: Emergence of resilience as a framework for state Departments of Transportation (DOTs) in the United States. Transp. Res. Part D: Transp. Environ. **82**, 102178 (2020)
5. Marsden, G., Anable, J., Shires, J., & Docherty, I. Travel Behaviour Response to Major Transport System Disruptions: Implications for Smarter Resilience Planning (International Transport Forum Discussion Paper 2016/09). (2016).
6. Mitoulis, S.-A., Bompa, D.V., Argyroudis, S.: Sustainability and climate resilience metrics and trade-offs in transport infrastructure asset recovery. Transportation Res. Part D: Transp. Environ. **121**, 103800 (2023)
7. Karagiannakis, G., Giordano, P.F., Limongelli, M.P.: Resilience of transport infrastructure against natural hazards: attributes and indicators. In: IABSE Congress San José 2024, pp. 698–706 (2024)
8. Ren, X., Zhu, D.: A conceptual framework for understanding the link between human and infrastructure resilience. Int. J. Disaster Risk Reduct. **50**, 101745 (2020)
9. Müllerklein, D., Fontaine, P.: Resilient transportation network design with disruption uncertainty and lead times. Eur. J. Oper. Res. **322**(3), 827–840 (2025)
10. Aultman-Hall, L., Dowds, J. Challenges and Opportunities for Integrating Climate Adaptation Efforts Across State, Regional and Local Transportation Agencies. (2015)
11. Svenson, P., Eriksson, K., Janhäll, S.: Resilience in Systems of Systems: Electrified Transport Systems, pp. 162–167 (2021)

ViBERT-ESG: Advancing ESG Classification in Vietnamese with ViEnESG-100 Dataset

Vanmanh Nguyen, Ductrung Tran, Kienlam Do, Thiloan Bui, and Vanha Tran(✉)

FPT University, Hanoi 155514, Vietnam
{manhnvhe170552,trungtdhe173146,lamdkhe172132}@fpt.edu.vn,
{loanbt7,hatv14}@fe.edu.vn

Abstract. Environmental, Social, and Governance (ESG) factors play a critical role in assessing corporate sustainability and ethical impact. While recent advances in natural language processing (NLP) have leveraged large pre-trained models for ESG classification, most prior work focuses on English-language data, limiting applicability in multilingual contexts. In this work, we present a BERT-based approach for ESG classification tailored to Vietnamese financial and sustainability reports. A key contribution of our work is the construction of the ViEn-ESG-100 dataset, a high-quality, manually labeled bilingual Vietnamese–English ESG dataset containing 103,656 sentences, which we plan to release publicly to foster further research. We fine-tune BERT-based models on this dataset and obtain strong performance across multiple ESG dimensions. Our results highlight the effectiveness of pre-trained language models for low-resource, domain-specific ESG classification and provide a useful evaluation resource for Vietnamese ESG text classification. Our dataset, model and code are available at https://github.com/nguyen599/ViBERT-ESG.

Keywords: ESG · Classification · BERT · Multilingual Dataset

1 Introduction

Environmental, Social, and Governance (ESG) factors have gained significant attention as crucial indicators for assessing corporate sustainability, ethical impact, and long-term value creation [8]. Evaluating a company's ESG performance reveals insights into non-financial aspects of their operations and helps stakeholders understand their contributions to sustainable development [20]. The increasing need for objective and efficient evaluation of companies' ESG activities has highlighted limitations in traditional manual assessment methods. Manual analysis of large volumes of ESG-related text, such as sustainability reports and news articles, is time-consuming and prone to human error and subjective judgment. In addition, differences in evaluation criteria across ESG rating agencies make consistent assessment difficult [10].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 226–235, 2027.
https://doi.org/10.1007/978-3-032-30539-8_16

Recent advances in NLP provide promising tools for automating ESG analysis. Existing studies have applied NLP to tasks such as ESG scoring, text classification, risk detection, and information extraction from reports and financial news [13].

A significant focus in this research area involves leveraging powerful pretrained language models (PLMs), particularly those based on the Transformer architecture like BERT [3] and its variants like RoBERTa [12], DistilBERT [17], Sentence-BERT [16], FinBERT [5]. When it comes to identifying language and retrieving pertinent data for ESG evaluation, these models have proven to perform quite well. For instance, the E-BERT model [20], specifically trained based on Google's BERT architecture for ESG rating tasks, achieved a notable 93% accuracy in automating the evaluation process, aiming for accurate and reliable results. In tasks such as recognizing ESG concerns or classifying words as sustainable or unsustainable, several BERT-based models have demonstrated great performance. Furthermore, SEC-BERT finetune [19] developed a domain-adapted SEC-BERT [13] model for detecting ESG-related issues from financial news, achieving improved performance through in-domain pre-fine-tuning. Their system classifies news into 33 ESG issue types and powers the ESG Issue Detector (EID) tool, enabling investors to identify ESG risks efficiently. The study also explored multilingual augmentation and LLM-based zero-shot baselines, finding domain-specific models to be significantly more effective.

ESG-KIBERT [9], a BERT-based model further pre-trained on ESG-specific corpora, which incorporates a hard attention mechanism to enhance ESG classification performance. The model achieved 99.72% accuracy on a four-class (E, S, G, None) task and was combined with sentiment analysis and industry-specific weighting using SASB's materiality map to generate transparent, domain-aware ESG ratings. ESGify [7], an open-source NLP-based model built on MPNet [18] for multi-label ESG risk classification. Leveraging a manually annotated dataset of 14,000 sentences from ESG reports, the model achieved superior performance to GPT-3.5 using tailored data augmentation strategies such as back translation and LLM-generated examples.

Despite these advances, one of the main challenges is the lack of comprehensive, labeled ESG datasets—especially for low-resource languages. While multilingual ESG tasks are emerging, involving datasets in languages like Chinese [14], English [1], French, Japanese [6], and Korean [11], a significant portion of the existing research has focused primarily on English-language data. This limits the applicability of existing ESG models in other languages, especially in low-resource settings where labeled data is limited.

To address this gap, we introduce a Vietnamese–English ESG classification pipeline based on a new bilingual dataset and a comparative study of transformer backbones. Beyond ESG classification, this work also contributes to multilingual NLP through bilingual dataset construction, annotation quality control, and cross-lingual evaluation. Our main contributions are as follows:

1. We introduce the **ViEn-ESG-100** dataset, a new large-scale, bilingual (Vietnamese–English) dataset comprising 103,656 sentence-level ESG samples.

Recognizing the general lack of such resources, we created this dataset to support and accelerate future research in Vietnamese ESG analysis and multilingual ESG text classification.

2. We present BERT-based ESG classification models tailored to Vietnamese financial and sustainability reports, and we provide a comparative evaluation across several transformer backbones in both Vietnamese and English settings.

3. This work provides strong baseline results for automated ESG classification in Vietnamese and offers practical resources to support ESG analysis for stakeholders and researchers.

2 Methodology

Our methodology is designed for the low resource setting of Vietnamese ESG analysis while taking advantage of recent progress in NLP. We fine tuning pre-trained BERT-based models for the ESG classification task, since these models have shown strong performance across many NLP applications. This approach is suitable for Vietnamese because large annotated ESG datasets are still limited. By adapting pre-trained models to our ESG dataset, we aim to build a classification system that is accurate, efficient, and able to capture the characteristics of ESG reporting in Vietnamese documents. The overall ESG classification pipeline is shown in Fig. 1.

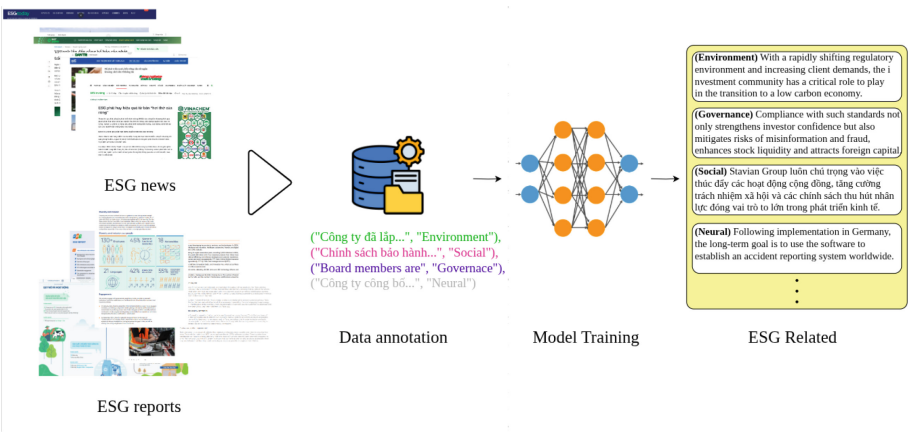


Fig. 1. Overview of the proposed ESG classification pipeline, including data preparation, bilingual annotation, model fine-tuning, and evaluation.

2.1 Model Selection

We experiment with several BERT-based backbones for ESG classification to compare how different pre-training strategies perform in our bilingual setting.

Our goal is not to propose a new architecture, but to examine which transformer models adapt most effectively to Vietnamese ESG text and whether similar trends also appear in English.

We focus on transformer models because they have become a standard choice for sentence classification, information extraction, and other NLP tasks where context matters. ESG statements often contain indirect wording, long dependencies, and domain-specific expressions, so models that can capture contextual meaning are a natural fit for this problem.

Our comparison includes BERT-base-cased [3], RoBERTa-base [12], XLM-RoBERTa-base [2], DeBERTa-v3-base and DeBERTa-v3-small [4], and DistilBERT-base-cased [17]. These models differ in size, pre-training strategy, and multilingual capacity, which makes them useful reference points for our experiments. BERT serves as a strong baseline, RoBERTa and DeBERTa represent alternative pre-training and encoding strategies, XLM-RoBERTa is included for its multilingual design, and DistilBERT provides a lighter-weight alternative with lower computational cost.

We fine-tune these models instead of training from scratch because our dataset is still much smaller than the large corpora typically used for language model pre-training. Using pre-trained checkpoints allows the models to reuse general language knowledge and focus on learning ESG-specific patterns. This approach is also more practical, since adapting existing models is more realistic than training a new model from scratch.

3 Dataset

The **ViEn-ESG-100** dataset is the main empirical resource used in this study. We created it to support ESG sentence classification in a bilingual Vietnamese–English setting, with particular attention to the lack of public Vietnamese data for this task. The final dataset contains 103,656 sentence-level samples and is intended to support both model training and controlled evaluation across ESG categories.

3.1 Data Collection

The sentences constituting the ViEn-ESG-100 dataset were sourced from a diverse range of publicly available documents relevant to corporate activities. The authors used two main data sources: ESG articles and reports. Articles were collected with the tag “ESG” from January 1, 2015 to January 14, 2025, for English articles, data were collected at ESGToday. Total of 78,291 sentences collected from the news. Annual ESG earnings report data from Vietnamese and S&P500 companies across ten key economic sectors, including technology, retail, consumer goods, health care, automotive, aviation, electronics, communication services, finance, and heavy industry, with publication dates spanning from 2012 to 2024 to capture evolving ESG practices and reporting standards over more than a decade. We collected 448 reports and a total of 45,103 text

data were manually collected. Table 1 shows the class-wise distribution of the ViEn-ESG-100 dataset.

Table 1. Class-wise distribution of the ViEn-ESG-100 Dataset.

Data		No. of Sentences
English		Vietnamese
Environment	10,517	14,712
Social	11,112	15,201
Government	9,103	11,030
Neural	15,210	16,771
Total	45,942	57,714

3.2 Annotation Methodology and Quality Assurance

To improve label quality, we relied on expert manual annotation rather than weak supervision or automatic labeling. Five trained annotators with backgrounds in finance and sustainability followed a shared guideline that defined the Environmental, Social, Governance, and Neutral categories. The work was carried out in two rounds: an initial annotation pass and a cross-check pass. Sentences that still lacked unanimous agreement after review were removed from the final dataset.

The Vietnamese portion contains 57,714 sentences and was labeled with attention to local reporting patterns and language use. The English portion contains 45,942 sentences drawn from public ESG disclosures and annual reports and was independently reviewed by three domain experts before adjudication. We measured agreement with Fleiss' Kappa, and the scores are reported in Table 2. Overall, the agreement values indicate that the annotation protocol was stable across both languages and suitable for downstream evaluation.

Table 2. Fleiss's Kappa coefficients measuring inter-annotator agreement across ESG categories and languages.

Category	English	Vietnamese
Environment	0.94	0.92
Social	0.91	0.90
Governance	0.93	0.89
Neutral	0.95	0.93
Overall	**0.93**	**0.91**

The agreement scores support the reliability of the annotation process and suggest that the label definitions are sufficiently clear for bilingual ESG classification. This manually curated dataset is therefore used as the core training

and evaluation resource throughout the paper. Figure 2 shows the token-length distribution obtained with the BERT-base-multilingual-cased tokenizer.

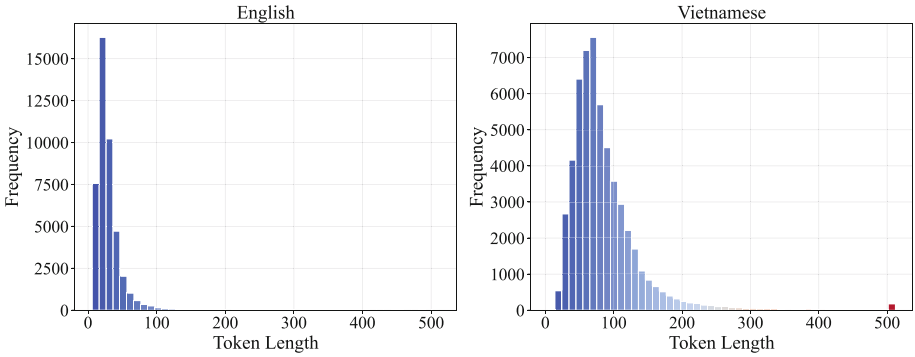


Fig. 2. Token-length distribution of the ViEn-ESG-100 dataset under the BERT-base-multilingual-cased tokenizer.

4 Results and Evaluation

This section describes the training setup and summarizes the main empirical results. We report performance for both English and Vietnamese and compare our bilingual models with several existing ESG baselines under binary and four-class settings.

4.1 Model Configuration and Training Strategy

We trained two types of models. The first group consists of three binary classifiers, one for each ESG pillar: Environmental, Social, and Governance. The second group consists of four-class models that assign each sentence to one of the labels E, S, G, or Neutral. The four-class setting is our main evaluation setup because it reflects the full label space of the dataset, while the binary models provide an additional view of category-specific performance.

Each experiment was conducted on a computer equipped with four RTX 3060 GPUs, 96 GB of RAM, and an Intel Core i7-13700 CPU. Torch 2.6.0, Datasets 3.6.0, Accelerate 1.7.0, and Transformers 4.51.0 are required for the implementation.

We used DeepSpeed ZeRO-3 optimization for full-parameter fine-tuning during training [15], 2,000 steps including 100 warm-up steps, were used to train the models in BFloat16. We used AdamW with a learning rate of 2×10^{-5}, cosine scheduling, weight decay of 0.01, a batch size of 48, a maximum gradient norm of 1, and a maximum input length of 512 tokens.

4.2 Experimental Result

Standard classification measures were used to assess the performance of our trained models, with an emphasis on the F1-score. The F1-score, which represents the harmonic mean of precision and recall, provides a balanced measure of a model's accuracy, particularly in scenarios with potential class imbalance, which can occur in ESG datasets where certain categories might be more prevalent than others. For the multi-class models, we offer F1-scores for each individual class (E, S, G, and N) as well as overall performance measures when applicable.

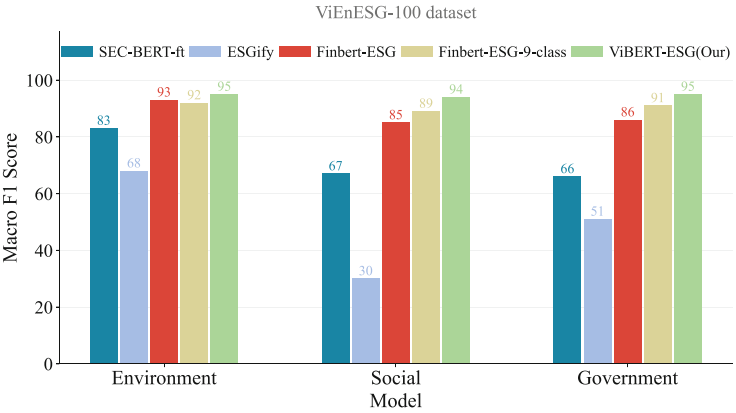


Fig. 3. Model Performance on ViEn-ESG-100 dataset.

For the English portion of the dataset, we benchmarked several publicly available ESG models, including FinBERT-ESG, SEC-BERT-ft, ESGify, and three binary classification models. These models differ in label granularity: FinBERT-ESG-9-class predicts 9 ESG-related categories, SEC-BERT-ft supports 33 fine-grained issue types, and ESGify handles 47 labels. To enable consistent comparison, we mapped all predicted labels into a unified 4-class format E/S/G/N. Figure 3 illustrates the performance of these models on the ViEn-ESG-100 dataset.

For English, ViDeBERTa-ESG-base achieved the highest F1-score on the Environment category (95.50%), while ViXLMRoBERTa-ESG-base slightly outperformed others on Governance (95.47%). Across all models, our multilingual fine-tuned versions consistently surpassed existing English-only ESG models, demonstrating the benefit of cross-lingual semantic transfer from bilingual training. In particular, lightweight models such as ViDistilBERT-ESG-base maintained competitive results (average F1 of 94.6%) while reducing computational cost.

For the Vietnamese subset, ViDeBERTa-ESG achieved the strongest performance overall, with F1-scores of 95.24% (Environment) and 93.18% (Governance), indicating robust understanding of domain-specific ESG contexts. Mean-

Table 3. F1 scores of models on each ESG category in the English ViEn-ESG-100 dataset.

Model	Backbone	# Param	Env	Soc	Gov	Neu
SEC-BERT-ft	SEC-BERT-base	109M	83.12	66.77	66.53	60.30
FinBERT-ESG	FinBERT	109M	92.67	84.90	86.25	87.26
FinBERT-ESG-9-class*	FinBERT	109M	92.16	89.01	91.35	86.89
ESGify	MPNet-base	109M	67.72	30.20	50.76	43.44
EnvironmentBERT	DistilRoBERTa	82M	92.15	–	–	92.76
SocialBERT	DistilRoBERTa	82M	–	76.81	–	81.23
GovernanceBERT	DistilRoBERTa	82M	–	–	64.46	80.06
ViBERT-ESG	BERT-base-multilingual-cased	168M	93.76	94.53	94.98	**94.15**
ViRoBERTa-ESG-base	RoBERTa-base	124M	95.43	94.06	95.01	91.32
ViXLMRoBERTa-ESG-base	XLM-RoBERTa-base	278M	95.00	95.00	**95.47**	92.19
ViDeBERTa-ESG-base	DeBERTa-v3-base	184M	**95.50**	94.49	94.81	91.48
ViDeBERTa-ESG-small	DeBERTa-v3-small	141M	94.55	94.85	94.58	90.19
ViDistilBERT-ESG-base	DistilBERT-base-multilingual-cased	135M	95.15	**95.19**	94.33	91.75
ViBERT-Env**	BERT-base-multilingual-cased	168M	94.62	–	–	92.13
ViBERT-Soc**	BERT-base-multilingual-cased	168M	–	94.86	–	92.22
ViBERT-Gov**	BERT-base-multilingual-cased	168M	–	–	93.47	93.82

Table 4. F1 scores of models on each ESG category in the Vietnamese ViEn-ESG-100 dataset.

Model	Backbone	# Param	Env	Soc	Gov	Neu
ViBERT-ESG	BERT-base-multilingual-cased	168M	93.50	89.73	91.77	**91.78**
ViRoBERTa-ESG	RoBERTa-base	124M	93.41	91.49	89.93	84.32
ViXLMRoBERTa-ESG	XLM-RoBERTa-base	278M	93.45	91.02	91.69	90.41
ViDeBERTa-ESG	DeBERTa-v3-base	184M	**95.24**	89.36	**93.18**	85.23
ViDeBERTa-small-ESG	DeBERTa-v3-small	141M	92.90	87.79	90.63	81.48
ViDistilBERT-ESG	DistilBERT-base-multilingual-cased	135M	93.87	**91.98**	90.63	87.17
ViBERT-Env*	BERT-base-multilingual-cased	168M	94.87	–	–	91.15
ViBERT-Soc*	BERT-base-multilingual-cased	168M	–	91.07	–	90.29
ViBERT-Gov*	BERT-base-multilingual-cased	168M	–	–	92.62	90.11

while, ViDistilBERT-ESG demonstrated the best performance in the Social category (91.98%), highlighting its efficiency and adaptability despite fewer parameters. The consistency of results across both languages shows that bilingual fine-tuning effectively preserves semantic alignment while enhancing model generalization.

Since existing ESG models are primarily designed for English and do not support Vietnamese input, our models are fine-tuned jointly on both English and Vietnamese samples from the ViEn-ESG-100 dataset. The models are better able to generalize across languages and take advantage of the shared semantic

field of ESG-related concepts thanks to this multilingual training arrangement. We assess our improved models using test sets in both Vietnamese and English.

This evaluation provides a clear comparison of different model backbones and configurations across ESG categories. The results in Tables 3 and 4 show strong performance across all ESG dimensions in both English and Vietnamese. Overall, these findings establish a strong multilingual benchmark for ESG classification in English and Vietnamese settings.

5 Conclusion and Future Work

This study focuses on building a practical ESG text classification system for Vietnamese, where public datasets and evaluation resources remain limited. To address this gap, we introduced the **ViEn-ESG-100** dataset and evaluated several transformer-based models in both Vietnamese and English. The results show that pre-trained language models can achieve strong ESG classification performance after fine-tuning on a carefully constructed bilingual dataset.

The main contribution of this work is the creation of a bilingual ESG resource together with baseline benchmarks for Vietnamese ESG analysis. We hope the dataset and experimental results can support future research on multilingual and low-resource ESG NLP tasks.

Future work may extend the dataset with additional ESG documents, explore more advanced tasks such as ESG entity recognition and issue extraction, and evaluate these models in real-world ESG monitoring systems. Overall, we hope this work provides a useful foundation for future Vietnamese ESG NLP research.

References

1. Chen, C., et al.: Multi-lingual ESG issue identification. In: Proceedings of the Fifth Workshop on Financial Technology and Natural Language Processing and the Second Multimodal AI For Financial Forecasting, pp. 111–115. Association for Computational Linguistics (2023). https://aclanthology.org/2023.finnlp-2.18/
2. Conneau, A., et al.: Unsupervised cross-lingual representation learning at scale. arXiv preprint arXiv:1911.02116 (2019). https://doi.org/10.48550/arXiv.1911.02116
3. Devlin, J., Chang, M., Lee, K., Toutanova, K.: BERT: pre-training of deep bidirectional transformers for language understanding. CoRR abs/1810.04805 (2018). https://doi.org/10.48550/arXiv.1810.04805
4. He, P., Gao, J., Chen, W.: DeBERTa: decoding-enhanced BERT with disentangled attention. arXiv preprint arXiv:2006.03654 (2020). https://doi.org/10.48550/arXiv.2006.03654
5. Huang, A., Wang, H., Yang, Y.: FinBERT: a large language model for extracting information from financial text. Contemp. Account. Res. **40**(2), 806–841 (2023). https://doi.org/10.1111/1911-3846.12832
6. Kao, Y., et al.: IMNTPU at ML-ESG-3: transformer language models for multi-lingual ESG impact type and duration classification. In: Proceedings of FinNLP@LREC-COLING 2024, pp. 228–233. Association for Computational Linguistics (2024). https://aclanthology.org/2024.finnlp-1.38/

7. Kazakov, A., et al.: ESGify: automated classification of environmental, social, and corporate governance risks. Dokl. Math. **108**(S2), S529–S540 (2023). https://doi.org/10.1134/S1064562423701673
8. Kiriu, T., Nozaki, M.: A text mining model to evaluate firms' ESG activities: an application for Japanese firms. Asia-Pacific Finan. Markets. **27**(4), 621–632 (2020). https://doi.org/10.1007/s10690-020-09309-1
9. Lee, H., Kim, J., Jung, H.: ESG-KIBERT: a new paradigm in ESG evaluation using NLP and industry-specific customization. Decis. Support Syst. **193**, 114440 (2025). https://doi.org/10.1016/j.dss.2025.114440
10. Lee, J., Kim, M.: ESG information extraction with cross-sectoral and multi-source adaptation based on domain-tuned language models. Expert Syst. Appl. **221**, 119726 (2023). https://doi.org/10.1016/j.eswa.2023.119726
11. Lee, J., Son, G., Kim, M.: ESG-KOR: a Korean dataset for ESG-related information extraction and practical use cases. In: Findings of the Association for Computational Linguistics: EMNLP 2024, pp. 6627–6643. (2024). https://doi.org/10.18653/v1/2024.findings-emnlp.387
12. Liu, Y., et al.: RoBERTa: a robustly optimized BERT pretraining approach. arXiv preprint arXiv:1907.11692 (2019). https://doi.org/10.48550/arXiv.1907.11692
13. Loukas, L., et al.: FiNER: financial numeric entity recognition for XBRL tagging. In: Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics (2022). https://doi.org/10.18653/v1/2022.acl-long.303
14. Pontes, E., Benjannet, M., Ming, L.: Leveraging BERT language models for multilingual ESG issue identification. arXiv preprint arXiv:2309.02189 (2023). https://doi.org/10.48550/arXiv.2309.02189
15. Rajbhandari, S., Rasley, J., Ruwase, O., He, Y.: Zero: memory optimizations toward training trillion parameter models. arXiv preprint arXiv:1910.02054 (2019). https://doi.org/10.48550/arXiv.1910.02054
16. Reimers, N., Gurevych, I.: Sentence-BERT: sentence embeddings using Siamese BERT-networks. arXiv preprint arXiv:1908.10084 (2019). https://doi.org/10.48550/arXiv.1908.10084
17. Sanh, V., Debut, L., Chaumond, J., Wolf, T.: DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. arXiv preprint arXiv:1910.01108 (2019). https://doi.org/10.48550/arXiv.1910.01108
18. Song, K., Tan, X., Qin, T., Lu, J., Liu, T.: MPNet: masked and permuted pre-training for language understanding. arXiv preprint arXiv:2004.09297 (2020). https://doi.org/10.48550/arXiv.2004.09297
19. Soni, P., Ghosh, S., Naskar, S.: Detecting issues related to environmental, social, and corporate governance using SEC-BERT. In: Data Science and Applications, pp. 329–338. Springer, Singapore (2024). https://doi.org/10.1007/978-981-99-7820-5_27
20. Zhang, M., Shen, Q., Zhao, Z., Wang, S., Huang, G.: Optimizing ESG reporting: Innovating with E-BERT models in natural language processing. Expert Syst. Appl. **265**, 125931 (2025). https://doi.org/10.1016/j.eswa.2024.125931

Machine Learning Assisted Two-Level Mixed Finite Element Method for the Darcy–Forchheimer Model

Johnson Daddy Audu[1,3(✉)] and Faisal Fairag[2]

[1] Department of Mathematics and Natural Sciences, Prince Mohammed Bin Fahd University, Alkhobar 31952, Saudi Arabia
jaudu@pmu.edu.sa

[2] Department of Mathematics, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia
ffairag@kfupm.edu.sa

[3] Center for Cognitive Intelligence & AI Systems, Prince Mohammed Bin Fahd University, Alkhobar 31952, Saudi Arabia

Abstract. We develop and analyse a PINN-assisted two-level mixed finite element method for the Darcy–Forchheimer equation. The coarse nonlinear solve of the classical two-level scheme is replaced by a physics-informed neural network (PINN) surrogate; the network output is projected onto the coarse finite element space and used as the linearization point for the fine-grid correction. Numerical results demonstrate that the proposed approach achieves accuracy comparable to the classical method, with nearly first-order convergence for both velocity and pressure. In multi-query settings where the trained network is reused across repeated solves with fixed physical parameters, the one-time training cost is offset by savings on repeated coarse solves.

Keywords: Darcy–Forchheimer equation · Mixed finite element method · Two-level algorithm · Physics-informed neural networks · Porous media

1 Introduction

Flows through porous media arise in numerous scientific and engineering applications, including petroleum reservoir simulation and groundwater hydrology [3,9]. In classical models, Darcy's law provides a linear relation between the pressure gradient and the velocity field. When the flow velocity is sufficiently large, however, inertial effects become significant and the linear model is no longer adequate [6,13]. The Darcy–Forchheimer equation, which augments Darcy's law with a quadratic drag term, is the standard nonlinear extension used in such regimes.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 236–246, 2027.
https://doi.org/10.1007/978-3-032-30539-8_17

Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with boundary $\partial\Omega$. The steady Darcy–Forchheimer system reads

$$\frac{\mu}{\rho}K^{-1}u + \frac{\beta}{\rho}|u|u + \nabla p = f \quad \text{in } \Omega, \tag{1.1}$$

$$\operatorname{div} u = b \quad \text{in } \Omega, \tag{1.2}$$

$$u \cdot n = g \quad \text{on } \partial\Omega, \tag{1.3}$$

where u and p denote velocity and pressure, K is the permeability tensor, and $\mu, \rho, \beta > 0$ are physical parameters. The nonlinear term $|u|u$ models inertial effects and complicates both the analysis and the numerical solution.

Mixed finite element methods have proved effective for Darcy–Forchheimer-type models, since they approximate velocity and pressure simultaneously and preserve the divergence constraint. Discretization strategies include mixed methods [7,14] and block-centered finite difference methods [11,12]. The resulting discrete systems are nonlinear, however, and these systems grow quickly with mesh refinement.

Two-level methods address this difficulty by solving the nonlinear problem on a coarse mesh and computing a linearized correction on a fine mesh [4]. For the Darcy–Forchheimer model, Fairag and Audu [5] showed that, under the mesh relation $h = O(H^2)$, the two-level scheme achieves the same asymptotic accuracy as a fully nonlinear fine-grid solve while requiring less computational work. The coarse nonlinear problem must still be solved iteratively, however, and this cost remains significant in large-scale or multi-query settings.

Physics-informed neural networks (PINNs) [8,10] offer an alternative by embedding the governing equations into a training loss function. Unlike Newton, the PINN training cost does not grow with mesh refinement, making the reuse argument stronger on larger problems. We propose replacing the coarse nonlinear solve in the two-level scheme with a PINN surrogate. The network is trained to minimize the residual of the Darcy–Forchheimer system, its output is projected onto the coarse finite element space, and the projected field serves as the linearization point for the classical fine-grid correction. We adopt the same P_0–P_1 mixed discretization and linearization strategy as in [5].

We develop a hybrid PINN–finite element framework in which the PINN replaces the coarse nonlinear solver while the fine-grid correction retains the stability of the classical mixed method. We establish a residual-to-error estimate (Theorem 3) bounding the PINN coarse error by a computable indicator η_H^{NN}, and prove global convergence of the full scheme (Theorem 4). Numerical experiments confirm that the PINN-assisted method matches the convergence rates of the classical two-level scheme, and a multi-query experiment shows the training cost is recovered after roughly 47 repeated solves at fixed physical parameters.

Sections 2–5 develop the theory and algorithm; numerical results are in Sect. 6 and conclusions in Sect. 7.

2 Variational Formulation and Analytical Properties

Assume $b \in L^{6/5}(\Omega)$ and $g \in L^{3/2}(\partial\Omega)$ satisfy the compatibility condition

$$\int_\Omega b\,dx = \int_{\partial\Omega} g\,d\sigma. \tag{2.1}$$

Define the function spaces

$$X := L^3(\Omega)^2, \qquad M := W^{1,3/2}(\Omega) \cap L^2_0(\Omega),$$

where $L^2_0(\Omega) = \{q \in L^2(\Omega) : \int_\Omega q\,dx = 0\}$. Since Ω is bounded and $3 > 2$, the embedding $L^3(\Omega) \hookrightarrow L^2(\Omega)$ is continuous, so the $L^2(\Omega)$ duality pairing is well-defined on X.

We define the bilinear form

$$b(v,q) := \int_\Omega \nabla q \cdot v\,dx, \qquad (v,q) \in X \times M, \tag{2.2}$$

and the nonlinear operator

$$\mathcal{A}(v) := \frac{\mu}{\rho} K^{-1} v + \frac{\beta}{\rho}|v|v, \qquad v \in X. \tag{2.3}$$

Let $\kappa_{\min} > 0$ denote the smallest eigenvalue of K. The operator $\mathcal{A}$ satisfies the following properties (see [2,7]):

(i) *Hemicontinuity*: $t \mapsto (\mathcal{A}(v+tw), z)$ is continuous on $[0,1]$ for all $v, w, z \in X$.
(ii) *Monotonicity*: $(\mathcal{A}(v) - \mathcal{A}(w), v - w) \geq \frac{\mu}{\rho}\kappa_{\min}\|v - w\|^2_{L^2}$ for all $v, w \in X$.
(iii) *Continuity*: $\|\mathcal{A}(v) - \mathcal{A}(w)\|_{L^{3/2}} \leq C(1 + \|v\|_{L^3} + \|w\|_{L^3})\|v - w\|_{L^2}$ for all $v, w \in X$.
(iv) *Coercivity*: $(\mathcal{A}(v), v) \geq \frac{\mu}{\rho}\kappa_{\min}\|v\|^2_{L^2}$ for all $v \in X$.
(v) *Inf–sup*: $\inf_{0\neq q\in M} \sup_{0\neq v\in X} \frac{b(v,q)}{\|v\|_{L^3}\|\nabla q\|_{L^{3/2}}} = 1$.

2.1 Variational Formulation

The weak formulation of (1.1)–(1.3) reads: find $(u,p) \in X \times M$ such that

$$(\mathcal{A}(u), \phi) + b(\phi, p) = (f, \phi), \qquad \forall\, \phi \in X, \tag{2.4}$$

$$b(u,q) = -\int_\Omega bq\,dx + \int_{\partial\Omega} gq\,ds, \qquad \forall\, q \in M. \tag{2.5}$$

Under the assumptions $f \in L^{3/2}(\Omega)^2$, $b \in L^{6/5}(\Omega)$, $g \in L^{3/2}(\partial\Omega)$, and the compatibility condition (2.1), properties (ii), (iv), and (v) above together with [7, Theorem 3] guarantee that problem (2.4)–(2.5) admits a unique solution $(u,p) \in X \times M$.

3 Regularization and Linearization

Since $u \mapsto |u|u$ is not differentiable at $u = 0$, we regularize: $|v|_\varepsilon := \sqrt{|v|^2 + \varepsilon^2}$, $F_\varepsilon(v) := |v|_\varepsilon v$, and

$$\mathcal{A}_\varepsilon(v) = \frac{\mu}{\rho} K^{-1} v + \frac{\beta}{\rho} F_\varepsilon(v). \tag{3.1}$$

For fixed $\varepsilon > 0$, F_ε is twice continuously differentiable with Fréchet derivative $DF_\varepsilon(v)w = |v|_\varepsilon w + (v \cdot w/|v|_\varepsilon)v$ and bounded first and second derivatives [5, Lemma 5.1]. Taylor's theorem gives $F_\varepsilon(v) = F_\varepsilon(w) + DF_\varepsilon(w)(v-w) + \mathcal{R}_\varepsilon(v,w)$, where $\|\mathcal{R}_\varepsilon(v,w)\|_{L^2} \le C\|v-w\|_{L^2}^2$. This expansion underpins the linearized fine-grid correction in Sect. 4.

4 Two-Level Discretization

Let $\mathcal{T}_H$ be a shape-regular triangulation of Ω with mesh size H, and let $\mathcal{T}_h$ be a conforming refinement with mesh size $h < H$, obtained by regular refinement of $\mathcal{T}_H$.

4.1 Finite Element Spaces

For $\ell \in \{H, h\}$ define

$$X_\ell := \{v_\ell \in L^3(\Omega)^2 : v_\ell|_K \in P_0(K)^2,\ \forall K \in \mathcal{T}_\ell\},$$
$$M_\ell := \{q_\ell \in C^0(\Omega) \cap W^{1,3/2}(\Omega) : q_\ell|_K \in P_1(K),\ \textstyle\int_\Omega q_\ell\, dx = 0\}.$$

Velocity is approximated by piecewise constants and pressure by continuous piecewise linears. By [14, Proposition 4.1], there exists $\beta_0 > 0$, independent of ℓ, such that

$$\inf_{0 \neq q_\ell \in M_\ell} \sup_{0 \neq v_\ell \in X_\ell} \frac{b(v_\ell, q_\ell)}{\|v_\ell\|_{L^3(\Omega)} \|\nabla q_\ell\|_{L^{3/2}(\Omega)}} \ge \beta_0, \qquad \ell \in \{H, h\}. \tag{4.1}$$

4.2 Classical Two-Level Method

Step 1 (Coarse nonlinear solve). Find $(u_H, p_H) \in X_H \times M_H$ such that

$$(\mathcal{A}(u_H), \phi_H) + b(\phi_H, p_H) = (f, \phi_H), \qquad \forall\, \phi_H \in X_H, \tag{4.2}$$
$$b(u_H, q_H) = -\int_\Omega b q_H\, dx + \int_{\partial\Omega} g q_H\, ds, \quad \forall\, q_H \in M_H. \tag{4.3}$$

The monotonicity and coercivity of $\mathcal{A}$, together with (4.1), guarantee a unique solution (u_H, p_H) (see [7, Theorem 3]).

Step 2 (Fine-Grid Linearized Correction). The coarse velocity u_H is prolonged to $\mathcal{T}_h$ (still denoted u_H) and used as the linearization point in the regularized formulation. Find $(u_h, p_h) \in X_h \times M_h$ such that

$$(D\mathcal{A}_\varepsilon(u_H)u_h, \phi_h) + b(\phi_h, p_h) = (f, \phi_h) + (D\mathcal{A}_\varepsilon(u_H)u_H - \mathcal{A}_\varepsilon(u_H), \phi_h), \quad \forall\, \phi_h \in X_h, \tag{4.4}$$

$$b(u_h, q_h) = -\int_\Omega b q_h \, dx + \int_{\partial\Omega} g q_h \, ds, \qquad \forall\, q_h \in M_h. \tag{4.5}$$

Coercivity of $D\mathcal{A}_\varepsilon(u_H)$ and the inf–sup condition (4.1) ensure unique solvability of the linear system (4.4)–(4.5) [5, Section 5.2].

4.3 Error Estimates for the Classical Scheme

The following estimates are drawn from [1,5].

Theorem 1. *Let (u_H, p_H) solve (4.2)–(4.3) and (u_h^{NL}, p_h^{NL}) the fully nonlinear fine-grid problem. There exist constants $C > 0$, independent of H and h, such that*

$$\|u - u_H\|_{L^2} + \|\nabla(p - p_H)\|_{L^{3/2}} \le CH, \tag{4.6}$$

$$\|u - u_h^{NL}\|_{L^2} + \|\nabla(p - p_h^{NL})\|_{L^{3/2}} \le Ch. \tag{4.7}$$

Theorem 2. (Two-level estimate [5]) *Let (u_h, p_h) be the two-level approximation from (4.4)–(4.5). There exists $C > 0$, independent of h and H, such that*

$$\|u - u_h\|_{L^2} + \|\nabla(p - p_h)\|_{L^{3/2}} \le C(\varepsilon h + h + H^2 + \varepsilon). \tag{4.8}$$

Under $h = H^2$, this matches the fully nonlinear fine-grid rate (4.7).

5 PINN-Assisted Two-Level Method

We extend the classical scheme by replacing the coarse nonlinear solve with a PINN surrogate. The fine-grid correction is unchanged.

5.1 Algorithm

Step 1 (PINN Coarse Surrogate). Let $(\hat{u}(\cdot;\Theta), \hat{p}(\cdot;\Theta))$ be neural network approximations of the velocity and pressure fields with weights Θ. Define the equation residuals

$$r_{\text{mom}} = \mathcal{A}(\hat{u}) + \nabla\hat{p} - f, \tag{5.1}$$

$$r_{\text{div}} = \operatorname{div}\hat{u} - b, \tag{5.2}$$

$$r_{\text{bc}} = \hat{u}\cdot n - g, \tag{5.3}$$

and the physics-informed loss

$$J(\Theta) = \omega_1\|r_{\text{mom}}\|^2_{L^2(\Omega)} + \omega_2\|r_{\text{div}}\|^2_{L^2(\Omega)} + \omega_3\|r_{\text{bc}}\|^2_{L^2(\partial\Omega)}. \tag{5.4}$$

After training to obtain Θ^*, set $\hat{u} = \hat{u}(\cdot;\Theta^*)$ and $\hat{p} = \hat{p}(\cdot;\Theta^*)$, and enforce zero mean pressure via $\hat{p} \leftarrow \hat{p} - |\Omega|^{-1}\int_\Omega \hat{p}\, dx$.

Step 2 (Projection onto the Coarse Finite Element Space). Let $\Pi_H^u : L^2(\Omega)^2 \to X_H$ and $\Pi_H^p : W^{1,3/2}(\Omega) \to M_H$ be stable L^2-projections. Define

$$u_H^{NN} = \Pi_H^u \hat{u}, \qquad p_H^{NN} = \Pi_H^p \hat{p}. \tag{5.5}$$

Step 3 (Surrogate Quality Check). Before proceeding to the fine-grid solve, we measure how well the projected PINN pair satisfies the discrete equations. Define the coarse residual functional

$$R_H^{NN}(\phi, q) = (f, \phi) - (\mathcal{A}(u_H^{NN}), \phi) - b(\phi, p_H^{NN}) - b(u_H^{NN}, q) - \int_\Omega bq\, dx + \int_{\partial\Omega} gq\, ds, \tag{5.6}$$

and its dual norm

$$\eta_H^{NN} = \sup_{\phi \in X,\ q \in M(\phi,q) \neq (0,0)} \frac{|R_H^{NN}(\phi, q)|}{\|\phi\|_{L^3(\Omega)} + \|\nabla q\|_{L^{3/2}(\Omega)}}. \tag{5.7}$$

The supremum in (5.7) is over an infinite-dimensional space and is not evaluated numerically in this work; η_H^{NN} instead serves as the theoretical quantity through which Theorem 3 certifies the surrogate.

Step 4 (Fine-Grid Linearized Correction). Solve the linearized system (4.4)–(4.5) with u_H replaced by u_H^{NN} to obtain $(u_h, p_h) \in X_h \times M_h$. By Theorem 3, a small η_H^{NN} ensures the PINN coarse error is small, and Theorem 4 then controls the resulting fine-mesh error via (5.9).

The complete procedure is summarized in Algorithm 1.

Algorithm 1. PINN-Assisted Two-Level Mixed Finite Element Method

1: **Input:** Domain Ω; mesh sizes H, h; parameters μ, ρ, β, K; data (f, b, g).
2: Train PINN: minimize $J(\Theta)$ in (5.4).
3: Project: $u_H^{NN} = \Pi_H^u \hat{u}$, $p_H^{NN} = \Pi_H^p \hat{p}$.
4: Solve the linearized fine-grid system (4.4)–(4.5) with u_H replaced by u_H^{NN}.
5: **Output:** (u_h, p_h).

Remark 1. By Theorem 3, η_H^{NN} bounds the coarse approximation error without the exact solution, so it can be inspected before the fine solve to gauge how much the result should be trusted. In the present scheme the fine solve always proceeds; an adaptive variant that triggers retraining when η_H^{NN} is large is left for future work.

Theorem 3. *Let $(u, p) \in X \times M$ solve (2.4)–(2.5), and let $(u_H^{NN}, p_H^{NN}) \in X_H \times M_H$ be defined by (5.5). Assume (4.1) holds. There exists $C > 0$, independent of H, such that*

$$\begin{aligned} &\|u - u_H^{NN}\|_{L^2(\Omega)} + \|\nabla(p - p_H^{NN})\|_{L^{3/2}(\Omega)} \\ &\quad \le C\Big(\inf_{v_H \in X_H} \|u - v_H\|_{L^2(\Omega)} + \inf_{q_H \in M_H} \|\nabla(p - q_H)\|_{L^{3/2}(\Omega)} + \eta_H^{NN} \Big). \end{aligned} \tag{5.8}$$

Proof. (Proof sketch) Set $e_u = u - u_H^{NN}$ and $e_p = p - p_H^{NN}$. Subtracting (5.6) from (2.4)–(2.5) gives $(\mathcal{A}(u)-\mathcal{A}(u_H^{NN}),\phi)+b(\phi,e_p)+b(e_u,q) = R_H^{NN}(\phi,q)$. Testing with $(e_u,-e_p)$ and using monotonicity (ii) yields $c_0\|e_u\|_{L^2}^2 \le \eta_H^{NN}(\|e_u\|_{L^3} + \|\nabla e_p\|_{L^{3/2}})$. The inf–sup condition (v) and continuity (iii) bound $\|\nabla e_p\|_{L^{3/2}}$, and combining gives $\|e_u\|_{L^2} + \|\nabla e_p\|_{L^{3/2}} \le C\eta_H^{NN}$. The full estimate (5.8) follows by the triangle inequality.

5.2 Global Convergence of the PINN-Assisted Scheme

Theorem 4. *Let $(u,p) \in X \times M$ be the exact solution of (2.4)–(2.5), and let (u_h,p_h) be the PINN-assisted two-level approximation. Let (u_h^{NL},p_h^{NL}) denote the fully nonlinear fine-grid solution and (u_h^{TL},p_h^{TL}) the classical two-level approximation linearized about u_H. There exists $C>0$, independent of h, H, and η_H^{NN}, such that*

$$\|u-u_h\|_{L^2(\Omega)} + \|\nabla(p-p_h)\|_{L^{3/2}(\Omega)} \le C\big(\varepsilon h + h + H^2 + \varepsilon + H + \eta_H^{NN}\big). \quad (5.9)$$

Proof. (Proof sketch) Decompose $u-u_h = (u-u_h^{NL})+(u_h^{NL}-u_h^{TL})+(u_h^{TL}-u_h)$ and bound each term. The first two terms are bounded by Theorems 1 and 2. For the third, subtracting the PINN-assisted system from the classical two-level system and using Lipschitz continuity of $D\mathcal{A}_\varepsilon$ gives $T_3 \le C\|u_H^{NN} - u_H\|_{L^2}$. The triangle inequality and Theorems 1 and 3 then yield $\|u_H^{NN} - u_H\|_{L^2} \le C(H+\eta_H^{NN})$, which proves (5.9).

Under $h = H^2$ and $\eta_H^{NN} = O(H)$, estimate (5.9) gives first-order convergence in h, matching the classical two-level rate.

6 Numerical Experiments

All computations use the manufactured exact solution

$$u(x,y) = \begin{pmatrix} \sin(\pi x)\sin(\pi y) \\ \cos(\pi x)\sin(\pi y) \end{pmatrix}, \qquad p(x,y) = \sin(\pi x)\cos(\pi y), \quad (6.1)$$

on $\Omega = (0,1)^2$. The source term f, divergence data b, and boundary data $g = u\cdot n$ are derived from (1.1)–(1.3) so that (6.1) satisfies the system exactly. Unless stated otherwise, $\mu = \rho = \beta = 1$, $K = I$, and $\varepsilon = 10^{-6}$.

The PINN uses a fully connected network with three hidden layers of 40 neurons and hyperbolic tangent activations. It is trained by minimizing $J(\Theta)$ in (5.4) with equal weights $\omega_1 = \omega_2 = \omega_3 = 1$, using the Adam optimizer at learning rate 10^{-3} for 1200 epochs, with 800 interior and 300 boundary collocation points per batch. All computations were carried out in MATLAB ONLINE R2025b.

Table 1. One-grid mixed FEM errors.

H	$\|u - u_H\|_{L^2}$	Rate	$\|\nabla(p - p_H)\|_{L^{3/2}}$	Rate
1/8	1.8646e−01	–	4.0750e−01	–
1/16	9.5546e−02	0.9646	2.0604e−01	0.9839
1/32	4.8157e−02	0.9884	1.0328e−01	0.9963
1/64	2.4138e−02	0.9964	5.1674e−02	0.9991
1/128	1.2069e−02	0.9980	2.5840e−02	0.9990

Table 2. Classical two-level method errors ($h = H^2$).

H	h	$\|u - u_h\|_{L^2}$	Rate_u	$\|\nabla(p - p_h)\|_{L^{3/2}}$	Rate_p
1/2	1/4	3.698553e−01	–	7.909967e−01	–
1/4	1/16	1.032088e−01	0.9207	2.216191e−01	0.9178
1/8	1/64	2.648289e−02	0.9812	5.704493e−02	0.9790
1/16	1/256	6.707819e−03	0.9906	1.442337e−02	0.9918

6.1 Numerical Results

Table 1 verifies first-order convergence of the base discretization. Both two-level methods under $h = H^2$ (Tables 2 and 3) achieve nearly first-order convergence in h, consistent with Theorems 2 and 4. Figure 1 confirms the slopes are indistinguishable.

6.2 Effect of the Forchheimer Number

To assess robustness under stronger nonlinearity, we fix $H = 1/8$, $h = 1/64$ and vary $\beta \in \{1, 5, 10, 20\}$, comparing both methods at each value. The results are reported in Table 4.

At $\beta = 1$ both methods agree closely. As β increases, the PINN pressure error grows, by a factor of roughly 13 between $\beta = 1$ and $\beta = 20$, against a factor of about 3 for the classical scheme. The benefit of pressure prolongation diminishes at higher β, since the quality of p_H^{NN} is bounded by that of the underlying surrogate.

Table 3. PINN-assisted two-level method errors ($h = H^2$).

H	h	$\|u - u_h\|_{L^2}$	Rate_u	$\|\nabla(p - p_h)\|_{L^{3/2}}$	Rate_p
1/2	1/4	3.459306e−01	–	7.830036e−01	–
1/4	1/16	9.560424e−02	0.9277	2.083135e−01	0.9551
1/8	1/64	2.429132e−02	0.9883	5.238011e−02	0.9958
1/16	1/256	6.267570e−03	0.9772	1.358593e−02	0.9735

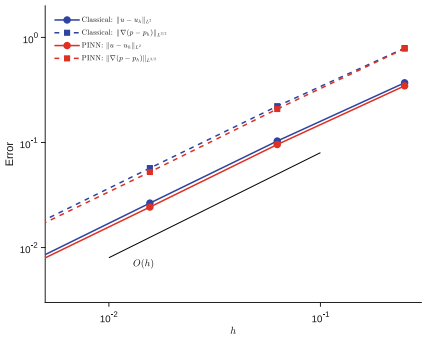


Fig. 1. Log–log convergence of both methods. Both follow the $O(h)$ slope.

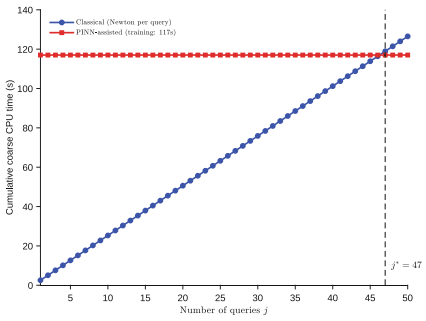


Fig. 2. Cumulative coarse CPU time at $H = 1/128$, $\beta = 20$. PINN becomes cheaper after $j^* = 47$ queries.

Table 4. Errors at fixed $H = 1/8$, $h = 1/64$ for varying β.

	Classical		PINN-assisted	
β	$\|u - u_h\|_{L^2}$	$\|\nabla(p - p_h)\|_{L^{3/2}}$	$\|u - u_h\|_{L^2}$	$\|\nabla(p - p_h)\|_{L^{3/2}}$
1	2.6483e−02	5.7045e−02	2.4373e−02	5.2486e−02
5	1.9430e−02	7.4865e−02	1.9212e−02	8.0660e−02
10	1.7977e−02	1.0034e−01	2.3777e−02	1.7758e−01
20	1.7304e−02	1.5820e−01	2.6188e−02	6.7849e−01

6.3 Computational Cost

Table 5 reports CPU times for both methods. PINN training costs roughly 26–28 s at every mesh size, while Newton on the coarse mesh takes 15–48 ms. The fine-grid times are close—both methods solve the same linearized system—though at the finest mesh the PINN fine solve is slightly faster (118s vs 124s). In a single run the PINN-assisted method is more expensive; Sect. 6.4 shows when this reverses.

6.4 Multi-query Reuse

To test the reuse argument, we fix $H = 1/128$, $\beta = 20$, and train the PINN once on the base manufactured solution. We then run 50 queries with varying source frequency $j = 1, 2, \ldots, 50$, where each query uses the manufactured solution

$$u^{(j)}(x, y) = \begin{pmatrix} \sin(j\pi x)\sin(\pi y) \\ \cos(j\pi x)\sin(\pi y) \end{pmatrix}, \quad p^{(j)}(x, y) = \sin(j\pi x)\cos(\pi y),$$

giving a different source term $f^{(j)}$ while keeping β, K, μ, ρ fixed. For each query, the classical method pays a full Newton solve on the coarse mesh; the

Table 5. CPU time comparison (seconds).

H	h	Method	Coarse (s)	Fine (s)	Total (s)
1/4	1/16	Classical	1.91e−02	4.34e−02	6.26e−02
1/4	1/16	PINN-assisted	2.76e+01	2.98e−02	2.76e+01
1/8	1/64	Classical	1.52e−02	1.57e+00	1.58e+00
1/8	1/64	PINN-assisted	2.62e+01	1.59e+00	2.78e+01
1/16	1/256	Classical	4.77e−02	1.24e+02	1.25e+02
1/16	1/256	PINN-assisted	2.67e+01	1.18e+02	1.45e+02

PINN method pays nothing at the coarse stage since the network is already trained. The fine-grid cost is identical for both methods and does not affect the comparison.

Training took 117 s, against an average Newton time of 2.53 s per query. The PINN cumulative cost crosses below the classical at query $j^* = 47$, in close agreement with the theoretical estimate $t_{\text{train}}/t_{\text{Newton}} \approx 46$. Figure 2 shows the cumulative coarse CPU time for both methods as a function of query number.

7 Conclusion

We introduced a PINN-assisted two-level mixed finite element method for the Darcy–Forchheimer equation. The PINN replaces the coarse nonlinear solve; its output is projected onto the coarse FE space and used to linearize the fine-grid problem. Theorem 3 bounds the PINN coarse error via η_H^{NN}, and Theorem 4 establishes first-order convergence under $h = H^2$. The method matches the classical scheme at $\beta = 1$; pressure errors grow at higher β. A multi-query experiment shows the PINN becomes cheaper after $j^* = 47$ repeated solves.

References

1. Audu, J.D., Fairag, F., Mustapha, K.: Mixed finite element analysis for generalized Darcy-Forchheimer model in porous media. J. Comput. Appl. Math. **353**, 191–203 (2019)
2. Audu, J.D., Fairag, F.A., Messaoudi, S.A.: On the well-posedness of generalized Darcy-Forchheimer equation. Bound. Value Probl., 123 (2018)
3. Bear, J.: Dynamics of Fluids in Porous Media. Elsevier (1972)
4. Fairag, F.: Two level finite element technique for pressure recovery from stream function formulation of the Navier-Stokes equations. In: Barth, T.J., Chan, T., Haimes, R. (eds.) Multiscale and Multiresolution Methods, pp. 297–306. Springer (2002)
5. Fairag, F., Audu, J.: Two-level Galerkin mixed finite element method for the Darcy-Forchheimer model in porous media. SIAM J. Numer. Anal. **59**(2), 888–911 (2021)
6. Forchheimer, P.: Wasserbewegung durch Boden. Zeitschrift des Vereines Deutscher Ingenieure **45**, 1782–1788 (1901)

7. Girault, V., Wheeler, M.F.: Numerical discretization of a Darcy-Forchheimer model. Numer. Math. **110**(2), 161–198 (2008)
8. Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., Yang, L.: Physics-informed machine learning. Nat. Rev. Phys. **3**, 422–440 (2021)
9. Muskat, M.: The Flow of Homogeneous Fluids Through Porous Media. McGraw-Hill (1946)
10. 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)
11. Rui, H., Pan, H.: A block-centered finite difference method for the Darcy-Forchheimer model. SIAM J. Numer. Anal. **50**(5), 2612–2631 (2012)
12. Rui, H., Zhao, D., Pan, H.: A block-centered finite difference method for Darcy-Forchheimer model with variable Forchheimer number. Numer. Methods Partial Differ. Equ. **31**, 1603–1622 (2015)
13. Ruth, D., Ma, H.: On the derivation of the Forchheimer equation by means of the averaging theorem. Transp. Porous Media **7**(3), 255–264 (1992)
14. Salas, J.J., López, H., Molina, B.: An analysis of a mixed finite element method for a Darcy-Forchheimer model. Math. Comput. Model. **57**(9), 2325–2338 (2013)

On the Travelling Wave Solutions of a Non-linear Reaction-Diffusion-Advection Model for Forest Fire Spread

Donato Pera[1(✉)], Cordula Reisch[2], Federica Di Michele[3], and Bruno Rubino[1]

[1] Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila, via Vetoio, loc. Coppito, I, L'Aquila, Italy
donato.pera@univaq.it

[2] Department of Science and Environment, Roskilde University, Universitetsvej 1, Roskilde, Denmark

[3] Istituto Nazionale di Geofisica e Vulcanologia, via Alfonso Corti 12, Milan, Italy

Abstract. This preliminary work investigates the dynamics of forest fire propagation using an Advection-Diffusion-Reaction (ADR) model with a non-linear diffusion term. A central focus of this work is the study of specific conditions under which travelling wave solutions occur within the ADR considered model, extending existing literature to a more general, one-dimensional (1D), non-linear diffusion case. The model approximation is validated through numerical simulations, demonstrating its effectiveness in capturing the front dynamics of wildfire propagation. The results confirm the model's reliability in identifying the critical wave speed of propagation, for the studied cases, offering a potential valuable tool for risk mitigation and strategic territorial planning.

Keywords: Forest fire model · advection-reaction-diffusion model · travelling waves

1 Introduction

Forest fires pose a significant threat to vulnerable areas due to the extensive damage they cause to both socio-economic activities and natural ecosystems. Consequently, it is essential to develop robust mathematical models and, subsequently, simulation software useful for estimating the dynamics of forest fires, possibly in real time. These tools are crucial for enhancing damage containment during firefighting operations and for the strategic planning e.g. of industrial sites near forested areas. Fast and reliable computational techniques are indispensable for effective risk mitigation. Therefore, the computational cost required for the simulation itself is crucial. Among the various models available in the literature, this work considers a model based on advection-diffusion-reaction (ADR)

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 247–258, 2027.
https://doi.org/10.1007/978-3-032-30539-8_18

equations, originally presented in [1] and recently investigated in [9,10]. In addition to computational techniques, it can be highly useful to study the triggering conditions of travelling wave solutions and to derive analytical estimates for the travelling wave speed. In this paper, we will derive an estimate of the travelling wave speed based on a general solution for ADR equations. The estimate will hold under the assumption of a constant biomass and relates to other estimates of the travelling wave speed under specific conditions, like in [1,5,8,10]. The set of time-evolution parabolic equations of reaction-diffusion type includes advection terms and non-linear diffusion. We compare the analytical predicted minimal travelling wave speed with a numerical solution computed with a spatial centered finite differences method and time integration through an explicit Euler scheme. The use of explicit methods and finite differences is limited by stability conditions on the time step. However, given the preliminary nature of the work and the analytical study carried out in 1D, no particular problems are encountered in this case from a computational point of view. The outline of the work is the following: Sect. 2 illustrates the model, Sect. 3 analyses the existence of travelling wave solutions and estimates the travelling wave speed. Section 4 reports numerical tests to validate the analytical results obtained in Sect. 3, and finally, Sect. 5 reflects the results and conclusions of the work.

2 Mathematical Model

This section provides a brief overview of mathematical models for forest fire dynamics. An extended and detailed taxonomy of the various approaches and models is provided in [2,3]. It is possible to group mathematical forest fire models into three categories such as physics-based models, data-driven models and mechanistic surrogate models. The data-driven models use empirical data to "identify" the fire dynamics in particular cases (sometimes using AI-based tools), whereas mechanistic surrogate models operate independently of the actual physical mechanics of the phenomenon. Finally, the theoretical models are based on the fundamental physical laws that describe the underlying processes. In this work, we utilize a physics-based theoretical model of advection-diffusion-reaction (ADR) type. The model describes the temperature evolution in a fuel-air-mixture during the fire spread in relation to the combustion of a biomass (fuel) fraction. The temperature is defined as a function $T(x,t)$, and the biomass fraction is described by the function $M(x,t)$. Both functions are defined over a one-dimensional domain $L \subset \mathbb{R}$, where the variable $x \in L$ represents a generic point within the domain and $t > 0$ denotes time. The formal derivation of the mathematical model is based on an energy balance, see for example [6,10]. The considered system of equations is

$$\begin{cases} \rho_0 C\left(\frac{\partial T}{\partial t} + \mathbf{c} \cdot \nabla T\right) & = \nabla \cdot (k_t(T) \cdot \nabla T) - h(T - T_\infty) + \psi(T)\rho_0 H M \\ \frac{\partial M}{\partial t} & = -\psi(T) M \end{cases} \tag{1}$$

where

$$\psi(T) = s(T)A\exp\left(-\frac{T_{\text{ac}}}{T}\right) \text{ with } s(T) = \begin{cases} 0, & \text{if } T < T_{\text{pc}} \\ 1, & \text{if } T \geq T_{\text{pc}} \end{cases}$$

is the combustion function, and $k_t(T) = 4\sigma\epsilon\delta T^3 + k$ gives a nonlinear diffusion term. The first equation of the system (1) has an advection term, a non linear diffusion term $k_t(T)$ where k is the thermal diffusivity, ϵ is an emissivity factor, δ is the optical path length for radiation, considered smaller than the characteristic length of the considered computational domain, and σ is Stefan-Boltzmann constant. The term $h(T - T_\infty)$ models Newton's cooling law related to convection and the environment temperature T_∞, and $\psi(T)\rho_0 HM$ models the combustion process in interaction with the biomass M, described by the second equation in (1). C is the specific heat and ρ_0 is the density of the air-fuel mixture described by the density of the fuel mass fraction. We define $T_{\text{ac}} = \frac{E_A}{R}$ with the activation energy E_A and the universal gas constant R, and T_{pc} is the ignition temperature. We examine two formulations for the advection term: the first incorporates a spatially constant velocity $v(t)$ to represent wind effects, while the second employs $c(x,t) = v(t) + \nabla z(x)$. From a physical standpoint, this latter configuration accounts for the influence of terrain topography on fire spread by integrating the elevation gradient. The topography gradient approach has been studied as well e.g. in [7].

3 Travelling Wave Solutions

In this section, we briefly recall the basic concepts related to travelling waves solution (TWS) for ADR equations. We do not claim to exhaustively explain all the theory relating to this type of solution, so please refer to [4,5,8,10] for further details. The basic idea for TWS for an ADR equation is to find a solution $u(x,t) = u(x + c_s t)$ where c_s is the velocity of the travelling wave and is, usually, unknown. Travelling wave solutions for the advection-reaction diffusion system (1) have been investigated with a linear diffusion term $k_t(T) = k$ and without Newton's cooling term in [8]. A linear approximations of the travelling wave speed, under the assumption of constant biomass and linear diffusion, was done in [10]. In [5], the nonlinear system in a persisting fire case has been transformed to an ODE system, which is a promising direction for proving the existence of TWS in this case. One target of this work is to define an estimate of the minimal velocity c^* for which the forest fire model (1) admits TWS. We will study the existence of TWS under the assumptions that:

1. we are in a setting with one spatial dimension;
2. the biomass can be considered constant in the time window relevant for the propagation of the wave front;
3. the fire is already ignited, by focusing the analysis on the characterization of the early-stage fire front dynamics.

Considering a one-dimensional model, we can rewrite equation (1) as

$$\begin{cases} \rho_0 C\Big(\frac{\partial T}{\partial t} + (v(t) + \frac{\partial Z}{\partial x})\frac{\partial T}{\partial x}\Big) & = \frac{\partial}{\partial x}(k_t(T)\frac{\partial T}{\partial x}) - h(T - T_\infty) + \psi(T)\rho_0 HM, \\ \frac{\partial M}{\partial t} & = -\psi(T)M. \end{cases} \tag{2}$$

Given the preliminary nature of this work, we will treat M as constant in order to study the behaviour of the ADR solutions over a short time scales, with the aim of understanding whether and how a travelling wave might be generated. As we assume M to be constant in time, we consider only the first equation in (2), which is a particular case of the more general equation

$$\left(\frac{\partial u}{\partial t} + \frac{\partial F(u)}{\partial x}\right) = \frac{\partial}{\partial x}\left(K(u)\frac{\partial u}{\partial x}\right) + R(u)\,, \tag{3}$$

where the function $u(x,t)$ will be a general function and in the same way we can define the functions: $F(u)$ as a general non-linear advection term, $K(u)$ as a general non-linear diffusion term and finally $R(u)$ as a general reaction term depending on the function $u(x,t)$. For Eq. (3) the existence of TWS has been investigated in [4]. For this class of equations, the following theorem holds:

Theorem 1 (Theorem 1.4 in [4]). *Assume that the functions $f = \frac{dF(u)}{du}$, $K(u)$ are continuous in the interval $[0,1]$, with $\min K(u) > 0$. If the Dini derivative $D^+(KR)(0) < +\infty$, then there exists a constant velocity c^* with*

$$2\sqrt{D_+(KR)(0)} - f(0) \le c^* \le 2\sqrt{\sup_{s\in]0,1]} \frac{K(s)R(s)}{s}} - \min_{s\in[0,1]} f(s) \tag{4}$$

such that Eq. (3) admits non-decreasing solutions if and only if $c > c^*$. Moreover, the solution is unique (up to translation of the origin). Whereas, if the Dini derivative $D_+(KR)(0) = +\infty$, Eq. (3) does not admit monotone solutions for any wave speed c.

The definition of Dini's derivatives is given for ease of reading and discussion:

Definition 1. *For the general continuous function $u(\cdot,t) : \mathbb{R} \to \mathbb{R}$ Dini's derivatives are*

$$D_+u(x,t) = \liminf_{h\to 0^+} \frac{u(x+h,t) - u(x,t)}{h},$$
$$D^+u(x,t) = \limsup_{h\to 0^+} \frac{u(x+h,t) - u(x,t)}{h}.$$

Non-dimensional Equation. According to Theorem 1, we need to analyze and rewrite our model in a non-dimensional way. As reported in [1], we use the transformations $\tilde{x} = \frac{x}{\rho_0}$, $\eta = \frac{M}{\rho_0}$, $\tau = \frac{t}{t_0}$, and $u = \frac{T - T_\infty}{\epsilon T_\infty}$, $w = \frac{t_0}{l_0}v$, where we also

have $\epsilon = \frac{RT_\infty}{E_A}$, $q = \frac{HM_0}{CT_\infty}$, $t_0 = \frac{e^{\frac{1}{\epsilon}}\epsilon}{qA}$, and finally $l_0 = \sqrt{\frac{t_0 k}{\rho_0 C}}$. We will drop the tilde for $\tilde{x}$ and write x instead. The non-dimensional equation then is

$$\frac{\partial u}{\partial t} + \left(w + \frac{\partial Z}{\partial x}\right)\frac{\partial u}{\partial x} = \frac{\partial}{\partial x}\left(K(u)\frac{\partial u}{\partial x}\right) + g(u,\beta)\,, \tag{5}$$

with $K(u) = \tilde{k}(1+\epsilon u)^3 + 1$, and $g(u,\beta) = s(u)^+ \beta e^{\frac{u}{1+\epsilon u}} - \alpha u$ where $\tilde{k} = \frac{4\sigma\delta T_\infty^3}{k}$ and $\alpha = \frac{t_0 h}{\rho_0 C}$. The function $s(u)^+$ is the part of the step function, where $s(u) = 1$, so for $u > 0$.

Case 1: Advection by Wind, Negligible Topographic Effects. In this section, we study the case where the advection term $F(u)$ is defined as $F(u) = wu$ with $w = w(t)$. Therefore, $\frac{\partial F}{\partial x} = \frac{\partial w}{\partial x}u + w\frac{\partial u}{\partial x} = w\frac{\partial u}{\partial x}$. We investigate the conditions required in Theorem 1 to have a non-decreasing solution of (5). The first condition is the continuity of the function $K(u)$ that is satisfied according to the definition. The second condition to verify is $\min_{u\in[0,1]} K(u) > 0$. According to the definition of the function $K(u)$ in Eq. (3), we have $\min_{u\in[0,1]} K(u) = \tilde{k} + 1 > 0$, because $u > 0$ in physical meaningful cases relevant for fire events where $T > T_\infty$. So, $\min K(u) > 0$ in the interval $[0,1]$. The second condition that we have to verify is related to the Dini's derivative $D^+(K(u)g(u,\beta))$, where we have

$$K(u)g(u,\beta) = \left(\tilde{k}(1+\epsilon u)^3 + 1\right)\left(s(u)^+ \beta e^{\frac{u}{1+\epsilon u}} - \alpha u\right).$$

According to the hypothesis to study an already ignited fire event, we can consider the term $s(u)^+ = 1$ and rewrite the previous equation, considering also the continuity of the reaction term:

$$K(u)g(u,\beta) = (\tilde{k}(1+\epsilon u)^3 + 1)(\beta e^{\frac{u}{1+\epsilon u}} - \alpha u).$$

We assume positivity of all parameters $\tilde{k}, \epsilon, \beta, \alpha$. According to Theorem 1 we have to check the condition $D^+ Kg(0) < +\infty$. For the sake of simplicity, we define the function $m(u) = K(u)g(u,\beta)$, so

$$m(u) = \left(\tilde{k}(1+\epsilon u)^3 + 1\right)\left(\beta e^{\frac{u}{1+\epsilon u}} - \alpha u\right).$$

We have to compute the Dini derivative $D^+ m(0)$ for $u = 0$ which is $D^+ m(u_0) = \lim_{h\to 0^+} \sup_{u\in[u_0,u_0+h]} \frac{(m(u)-m(u_0))}{h}$, we can compute $m(0) = (\tilde{k}+1)\cdot\beta$. The second step is to compute the behavior of the function $m(u)$ when $u \to 0^+$. We analyze the term $\tilde{k}(1+\epsilon\cdot u)^3 + 1$ for $u \to 0$, where for $(1+\epsilon\cdot u)^3$ using the Taylor formula we get $(1+\epsilon u)^3 = 1 + 3\epsilon u + \mathcal{O}(u^2)$. So we can rewrite the previous term as

$$\tilde{k}(1+\epsilon u)^3 + 1 = \tilde{k}(1 + 3\epsilon u + \mathcal{O}(u^2)) + 1. \tag{6}$$

Following, we can analyze the reaction term $\beta e^{\frac{u}{1+\epsilon u}} - \alpha u$. Using again the Taylor formula for the exponential term $e^{\frac{u}{1+\epsilon u}}$ per $u \to 0$, we get $\frac{u}{1+\epsilon u} =$

$u\left(1-\epsilon u+\mathcal{O}(u^2)\right)$. So, we can write

$$\beta e^{\frac{u}{1+\epsilon u}}-\alpha u=\beta+(\beta-\alpha)u+\beta(\frac{1}{2}-\epsilon)u^2+\mathcal{O}(u^3). \tag{7}$$

Combining the results in (6) and (7), we get

$$m(u)\approx(\tilde{k}+1+3\tilde{k}\epsilon u+\mathcal{O}(u^2))\cdot(\beta+(\beta-\alpha)u+\beta(\frac{1}{2}-\epsilon)u^2+\mathcal{O}(u^3))$$

and therefore $m(u)\approx(\tilde{k}+1)\beta+\left[(\tilde{k}+1)(\beta-\alpha)+3\tilde{k}\epsilon\beta\right]u+\mathcal{O}(u^2)$.

We can now compute the required Dini's derivative:

$$D^+m(0)=\left[(\tilde{k}+1)(\beta-\alpha)+3\tilde{k}\epsilon\beta\right],$$

and due to the positivity of the parameters $\tilde{k}$, ϵ, β and α, we have $D^+m(0)<\infty$. We now consider equation (4) to compute the minimal velocity c^*. First, we have to evaluate the term $\sqrt{\sup_{s\in(0,1]}\frac{K(s)g(s)}{s}}$. Therefore, we can define the function $l(s)=\frac{m(s)}{s}$ with $s\in(0,1]$, and $\tilde{k},\epsilon,\beta,\alpha>0$.

We can now compute $u\to 0^+$ using the Taylor formula near $u=0$ for $m(u)$ in Eq. (3), giving $m(u)\approx(\tilde{k}+1)\beta+\left[(\tilde{k}+1)(\beta-\alpha)+3\tilde{k}\epsilon\beta\right]u+\mathcal{O}(u^2)$ and then

$$l(s)=\frac{m(s)}{s}\approx\frac{(\tilde{k}+1)\beta}{s}+\left[(\tilde{k}+1)(\beta-\alpha)+3\tilde{k}\epsilon\beta\right]+\mathcal{O}(s).$$

This gives $\lim_{s\to 0^+}l(s)=+\infty$. We now compute $l(s)$ in $s=1$ and so we can write:

$$l(1)=m(1)=\left(\tilde{k}(1+\epsilon)^3+1\right)\left(\beta e^{\frac{1}{1+\epsilon}}-\alpha\right).$$

This value is a finite value depending on the parameters $\tilde{k},\epsilon,\beta,\alpha$. We can write now that $l(s)\to+\infty$ for $s\to 0^+$ and $l(1)$ is a finite value, so the supremum is $\sup_{u\in(0,1]}l(u)=+\infty$. To compute the c^* values, we have to consider again the function $m(u)$ and compute for the function the Dini derivative

$$D_+m(0)=\liminf_{h\to 0^+}\frac{m(h)-m(0)}{h},$$

where $m(0)=(\tilde{k}+1)\beta$. Considering now $u\to 0^+$, we get: $D_+m(0)=(\tilde{k}+1)(\beta-\alpha)+3\tilde{k}\epsilon\beta$. Recalling that $f=\frac{dF}{du}=w$ considering now w constant and so $f(0)=w$ and $\min_{s\in[0,1]}f(s)=w$. So finally we can write with Eq. (4)

$$c^*\geq 2\sqrt{(\tilde{k}+1)(\beta-\alpha)+3\tilde{k}\epsilon\beta}-w. \tag{8}$$

Obviously, to have the physical and mathematical meaning, it is required to have $(\beta-\alpha)\geq 0$ and that $2\sqrt{(\tilde{k}+1)(\beta-\alpha)+3\tilde{k}\epsilon\beta}\geq w$.

Case 2: Advection by Wind and Topography. In this section, we consider the case where $F(u) = \left(w + \frac{\partial Z}{\partial x}\right) \frac{\partial u}{\partial x}$ with a linear function $Z(x) = ax$ defined by $\frac{dZ(x)}{dx} = a$ with $a \in \mathbb{R}$ constant. Depending on the sign of the parameter $a \geq 0$ (or $a \leq 0$), the function Z represents an up-hill ramp or a down-hill ramp. We study these two cases in a similar way that we followed for the no topography case, indeed the advection term will still be linear with a modified function f. For applying Theorem 1, we have to check that $\min K(u) > 0$ in the interval $[0, 1]$. This condition will be satisfied according to the definition of the function $K(u)$ and the hypothesis of the problem. Indeed, we suppose to work with fire events already ignited and this hypothesis leads to having a temperature $T(x) > T_\infty$ and thus a function $u(x) > 0$ in the considered interval. Again, all requirements of Theorem 1 are fulfilled, and therefore we get estimated for the minimal spreading speed depending on the sign of a. For the up-hill case $a > 0$ we have $c^* \geq 2\sqrt{(\tilde{k}+1)(\beta-\alpha) + 3\tilde{k}\epsilon\beta} - (w + a)$, and for the down-hill case, we find $c^* \geq 2\sqrt{(\tilde{k}+1)(\beta-\alpha) + 3\tilde{k}\epsilon\beta} - (w - a)$.

Table 1. Simulation parameters with applicable range values and measure units for the considered ADR model as reported in [10]. In the last column have been reported the values used for the numerical simulations.

Parameter	Symbol	Values [10]	Measure Units	Simulation Values
Density	ρ_0	40–640	$kg \cdot m^{-1}$	40.0
Specific heat	C	1.0–3.0	$kJ \cdot kg^{-1} \cdot K^{-1}$	1.0
Pre-exponential factor	A	0.05	s^{-1}	0.05
Heating value	H	10–10000	$kJ \cdot kg^{-1}$	10000
Convection coefficient	h	0.1–10	$kW \cdot m^{-3} \cdot K^{-1}$	5
Ambient air temperature	T_∞	280–320	K	300
Ingnition temperature	T_{pc}	400–500	K	400
Activation temperature	T_{ac}	400–500	K	400
Thermal diffusivity	k	0–4	$kW \cdot m^{-1} \cdot K^{-1}$	4

4 Numerical Method, Simulations and Results

In this section, we introduce a discretization of the studied ADR model, for a dimensional case in 1D, using a finite difference method (FDM) scheme. The problem is solved using a FDM on a uniform spatial grid with step width Δx and an explicit time integration (Forward Euler) with step size Δt. In the numerical simulations we considered a general constant transport term named as v with a first-order upwind scheme: $v\frac{\partial T}{\partial x} \approx v\frac{T_j^n - T_{j-1}^n}{\Delta x}$. The diffusive term with variable conductivity $k(T)$ has been discretized by evaluating the heat fluxes at the cell interfaces $(j \pm 1/2)$: $\frac{\partial}{\partial x}\left[k(T)\frac{\partial T}{\partial x}\right] \approx \frac{1}{\Delta x}\left[k_{j+1/2}\frac{T_{j+1}^n - T_j^n}{\Delta x} - k_{j-1/2}\frac{T_j^n - T_{j-1}^n}{\Delta x}\right]$,

where the interface conductivities are calculated using the arithmetic mean: $k_{j+1/2} = \frac{1}{2}(k_j + k_{j+1})$ and $k_{j-1/2} = \frac{1}{2}(k_j + k_{j-1})$. The final discretized equation for updating the temperature at computational step time $n+1$ for each internal node named with j as counting variable is:

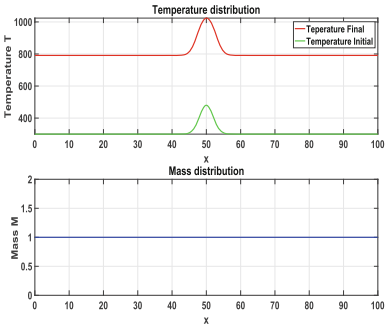
$$T_j^{n+1} = T_j^n + \frac{\Delta t}{\rho_0 C \Delta x^2}\left[k_{j+1/2}(T_{j+1}^n - T_j^n) - k_{j-1/2}(T_j^n - T_{j-1}^n)\right] - \frac{v\Delta t}{\Delta x}(T_j^n - T_{j-1}^n) - \frac{h\Delta t}{\rho_0 C}(T_j^n - T_\infty) + \frac{\Delta t}{C} A e^{-T_{\text{ac}}/T_j^n} H M_j. \quad (9)$$

The accuracy and stability of the numerical solution depends on the choice of spatial resolution, time step, and the physical properties. Since an explicit Forward Euler scheme is used, the time step Δt must satisfy the stability requirement for the diffusion term. Given the non-linear thermal diffusivity $\alpha(T) = k(T)/(\rho_0 C)$, the stability condition is defined as: $\Delta t \leq \min\left\{\frac{\Delta x}{\max(|v|)}, \frac{\rho_0 C \Delta x^2}{2\max(k(T))}\right\}$. The simulation is initialized with a Gaussian temperature profile with the hypothesis to study the front propagation when the ignition process already happened:

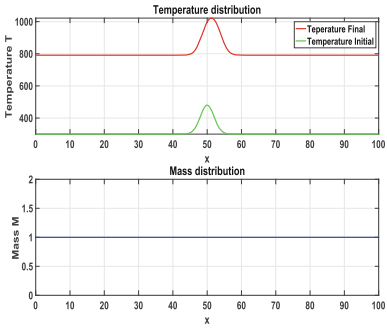
$$T(x,0) = T_{\text{base}} + (T_{\text{peak}} - T_{\text{base}}) \exp\left(-\frac{(x-x_0)^2}{2\sigma_g^2}\right), \quad (10)$$

where $T_{\text{peak}} = 480\,\text{K}$, $T_{\text{base}} = 300\,\text{K}$, $x_0 = L/2$, and $\sigma_g = 0.02L$. We considered a linear computational domain with length $L = 100$. Finally adiabatic boundary conditions (zero-gradient) have been applied in $x = 0$ and $x = L$ with $\frac{\partial T}{\partial x}\big|_{x=0} = 0$, $\frac{\partial T}{\partial x}\big|_{x=L} = 0$. In the discrete domain, these conditions have been implemented by equating the temperature of the boundary nodes to their immediate internal neighbors at each time step: $T_1^{n+1} = T_2^{n+1}$ and $T_{N_x}^{n+1} = T_{N_x-1}^{n+1}$. These conditions simulate an isolated system at the boundaries, ensuring that no heat flux enters or leaves the domain through conduction as also reported in [10]. As mentioned above, the computational domain has a physical length of $L = 100.0$ length units, and has been discretized into $N_x = 1000$ nodes. This choice leads to a spatial resolution of $\Delta x = 0.10$. The temporal evolution is integrated with a fixed time step of $\Delta t = 10^{-4}\,\text{s}$ over a total simulation time duration of 2.5 s. Finally we considered the Stefan-Boltzmann constant σ and scaled parameters $\epsilon = 1.0$ and $\delta = 0.09$. The other parameters used for the numerical simulations are summarized in Table 1. Specifically, four numerical simulations were conducted to evaluate TWS sensitivity with respect to the minimal velocity. The minimal velocity values have been computed for each test case, considering the dimensional effects, and denoted as c^{*d}. As reported in Fig. 1 we executed the four tests varying the advection velocities v with the aim to check the cases $v < c^{*d}$ and $v > c^{*d}$. Particularly, we considered the effects of the biomass using a constant or variable values for M, and v velocity equal to 0.030 or 0.5000 m/s with c^{*d} equal to 0.5388 and 0.0688 m/s respectively. About M in the constant cases, we considered M equal to 1.0 with $\frac{\partial M}{\partial t} = 0$, and in the variable case we started from M equal to 1.0 following the dynamics reported in equation (1). From the test results it is possible to see that for the cases 1(a) and 1(c), with

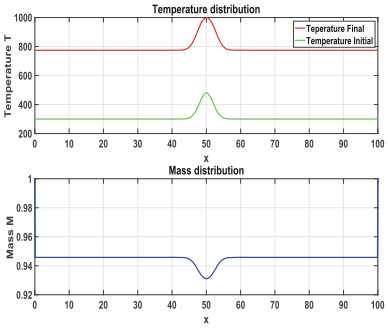
an analytical velocity limit equal to 0.5388 m/s, we have a solution almost fixed in the starting point without TWS as expected. In the other two cases 1(b) and 1(d), according to the analytical condition equal to 0.0688 m/s, we have a TWS that moves in the right direction with a numerical velocity comparable to the estimated limit analytical velocity.



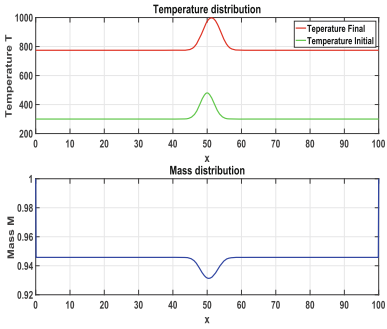
(a) M constant, $w = 0.03$ (m/s)



(b) M constant, $w = 0.50$ (m/s)



(c) M variable, $w = 0.03$ (m/s)



(d) M variable, $w = 0.50$ (m/s)

Fig. 1. Test cases with low (0.03 m/s) and high (0.50 m/s) advection velocity v and with M constant or variable for a simulation time equal to 2.5 s. Tests (a) and (c) show that the main front is almost stationary in the initial position, whereas in the cases (b) and (d) the main front moves according to the direction of advection term.

Numerical Results Analysis

In the considered non-linear models, the wave profile of $T(x,t)$ is subject to simultaneous phenomena: advection, diffusion and reaction. Figure 1 shows how the Gaussian profile used as initial data changes over time so tracking the wave-front velocity using a single point, such as the peak position $T_{\max}(t)$, could provides only a local description. Indeed the wave front may be subject to asymmetric deformations and the peak velocity could deviate from the rate at which

energy is transferred. For this reason, to describe the global propagation velocity, we also treated the numerical solutions as a spatial energy distributions, employing the center of mass ($X_{cm}(t)$) to compute the velocity of the 'mean position' of the heat across the studied cases

$$X_{cm}(t) = \frac{\int_0^L x \cdot [T(x,t) - T_\infty]\,\mathrm{d}x}{\int_0^L [T(x,t) - T_\infty]\,\mathrm{d}x}. \tag{11}$$

Finally the numerical velocity has been derived from the center of mass, $V_{cm} = \Delta X_{cm}/\Delta t$. In the light of application, the wave speed of the wave front has an high relevance. Therefore, we reported in Table 2 both the centre-of-mass velocity (following V_{cm}), which provides an estimate of the average speed at which the heat moves within the system, and the velocity of the maximum value $T_{\max}(t)$ named V_{fmax}. In Table 2 we reported also the used test velocities v and the limit velocities c^{*d} computed for the test cases described above. The numerical results show in the tested cases a good agreement with theory: in Table 2 we show the results of the four tests named (a), (b), (c) and (d) reporting for each test the biomass M as constant or variable, and the c^{*d}, v, V_{cm} values. Particularly, in the cases (a) and (c) we got a general good agreement between the theoretical wave behavior and the numerical tests with some almost stable solutions around the position of the initial values. Indeed, for these cases we used a v value that is smaller than the c^{*d} limit to generate TWS. In these two cases, for the sake of completeness, we computed the V_{cm} finding values close to zero and $V_{fmax} < v$ as expected (see Table 2). Conversely, in the cases (b) and (d) we got as expected a TWS with $V_{cm} > v$ and $V_{fmax} \approx v$ (see Table 2). Further, the comparison between the test cases with constant and variable biomass for the same wind velocity shows that the approximation of a constant biomass, used in the analytical derivation using Theorem 1, is reasonable in the studied time framework.

Table 2. Summary of the results related to the four tests (a), (b), (c) and (d) ref. Figure 1. Velocities V_{cm}, V_{fmax}, c^{*d} and v in m/s .

Test case	M	V_{cm}	V_{fmax}	c^{*d}	v
(a)	constant	0.0001	0.0015	0.5388	0.030
(b)	constant	0.5313	0.5009	0.0688	0.500
(c)	variable	0.0001	0.0015	0.5388	0.030
(d)	variable	0.5313	0.5009	0.0688	0.500

5 Conclusions and Future Works

This preliminary study explores the existence of travelling wave solutions for a non-linear advection-reaction-diffusion (ADR) system used in forest fire modeling. Specifically, following Theorem 1, we reduced the ADR system to a broader

class of equations and we verified that its properties satisfy the necessary theoretical requirements for the existence of travelling wave. We established the conditions for the existence of travelling wave solutions. Furthermore, numerical tests were performed to verify and support the analytical existence conditions. From the numerical results, we got a good agreement between the analytical and the numerical tests in the studied cases. Future research will focus on analyzing the problem under more complex topographical features, incorporating higher-order derivatives and extending the model to a two-dimensional case. Particularly, we will investigate a more general case study with a polynomial function for the effects, $Z(x) = ax^n$, with a first derivative with respect to x equal to $\frac{dZ(x)}{dx} = anx^{n-1}$. According to Eq. (5), it will be required to recompute the formula for the minimal wave speed c^* considering the new general advection term. Finally, further studies will conduct a sensitivity analysis with respect to all physical parameters that characterize the model and perform some comparisons with experimental cases insofar as useful data are reported in the literature.

Funding and Acknowledgments. This work was supported by European Union - NextGenerationEU under the Italian Ministry of University and Research (MUR) National Centre for HPC, Big Data and Quantum Computing CN_00000013-CUP:E13C22001000006.

References

1. Asensio, M.I., Ferragut, L.: On a wildland fire model with radiation. Numer. Meth. Eng. **54**(1), 137–157 (2002)
2. Grasso, P., Innocente, M.: Physics-based model of wildfire propagation towards faster-than-real-time simulations. Comput. Math. Appl. **80**(5), 790–808 (2020)
3. Innocente, M., Grasso, P.: Self-organising swarms of firefighting drones: Harnessing the power of collective intelligence in decentralised multi-robot systems. J. Comput. Sci. **34**, 80–101 (2019)
4. Malaguti, L., Marcelli, C.: Travelling wavefronts in reaction–diffusion equations with convection effects and non–regular terms. Math. Nachr. **242**, 148–164 (2002)
5. Mitra, K., Peng, Q., Reisch, C.: Studying wildfire fronts using advection–diffusion–reaction models. In: Sequeira, A., Silvestre, A., Valtchev, S.S., Janela, J. (eds.) Numerical Mathematics and Advanced Applications ENUMATH 2023, Volume 2, vol. 154, pp. 182–192, Cham (2025)
6. Navas-Montilla, A., Reisch, C., Diaz, P., Özgen Xian, I.: Modeling wildfire dynamics through a physics-based approach incorporating fuel moisture and landscape heterogeneity. Environ. Model. Softw. **192** (2025)
7. Nieding, L., Reisch, C., Langemann, D., Navas-Montilla, A.: Impact of topography and combustion functions on fire front propagation in an advection-diffusion-reaction model for wildfires. IFAC-PapersOnLine **59**(1), 103–108 (2025)
8. Babak, P., Bourlioux, A., Hillen, T.: The effect of wind on the propagation of an idealized forest fire. SIAM J. Appl. Math. **70**, 1364–88 (2009)
9. Pera, D., Michele, F.D., Stagnini, E., Rubino, B.: Analysis of Fonte Vetica fire (Abruzzo, Italy 2017) using an advection-reaction-diffusion model and

high-performance computing based numerical simulations. MEDGU24, Barcelona (Spain) (2024). Accepted for publication
10. Reisch, C., Navas-Montilla, A., Özgen Xian, I.: Analytical and numerical insights into wildfire dynamics: exploring the advection–diffusion–reaction model. Comput. Math. Appl. **158**, 179–198 (2024)

Anomaly Detection and Root Cause Analysis of Retainability Contextual Anomalies in Radio Access Networks

David Santos[1(✉)], Paula Amaral[1,2], João Martins[3], and Rui Ramos[3]

[1] NOVA School of Science and Technology, University Nova de Lisboa, Lisbon, Portugal
dsm.santos@campus.fct.unl.pt
[2] NOVA Math CMA, University Nova de Lisboa, Lisbon, Portugal
[3] Multivision, Lisbon, Portugal

Abstract. The increasing complexity of mobile communication networks, driven by the rapid growth of connected devices and data traffic, has significantly challenged network monitoring and troubleshooting for Mobile Network Operators. Traditional fault diagnosis approaches rely heavily on manual analysis, which is time-consuming, error-prone, and difficult to scale. This work proposes an automated framework for anomaly detection and root cause analysis in Radio Access Networks (RAN) using Machine Learning techniques. The approach integrates unsupervised anomaly detection methods with explainability tools to identify abnormal patterns in Key Performance Indicators (KPIs) and uncover their underlying causes. Dimensionality reduction and clustering techniques are employed to reveal intrinsic data structures, while model interpretability methods, such as SHAP, provide insights into feature importance and the decision-making process. Initial experimental results are promising. The proposed methodology shows potential in detecting anomalous behavior and provides preliminary support for root cause identification, although this aspect is still under development. Ongoing work focuses on refining the identification of underlying causes, with current results indicating a positive trajectory. This progress is expected to contribute to reducing troubleshooting time, improving operational efficiency, and enhancing overall network reliability.

Keywords: Root Cause Analysis · Machine Learning · Radio Access Networks · SHAP

1 Introduction

The widespread adoption of mobile devices has driven a significant expansion of mobile communication infrastructures to support an increasingly large and demanding user base. To ensure reliable operation, Mobile Network Operators (MNOs) must continuously monitor network performance and promptly diagnose potential faults.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 259–269, 2027.
https://doi.org/10.1007/978-3-032-30539-8_19

Despite this need, a substantial portion of the diagnostic process relies on manual investigation, with only limited automation available. As a result, troubleshooting remains a slow and resource-intensive task. The time required to resolve network issues can vary considerably depending on their complexity, often requiring engineers to spend hours analyzing performance data before identifying the root cause.

In this context, the development of automated tools to assist in network analysis becomes essential. Such tools can streamline the troubleshooting workflow, enabling engineers to handle a larger number of incidents in less time, while improving the overall efficiency and performance of the network.

Recent advances in Machine Learning (ML) have created new opportunities for automating network monitoring tasks, particularly in the areas of Anomaly Detection (AD) and Root Cause Analysis (RCA). These techniques allow for the detection of anomalous patterns within large volumes of performance data and provide insights into the factors contributing to network degradation.

In this paper, we present the progress achieved so far towards a data-driven approach for anomaly detection and root cause analysis in Key Performance Indicators (KPIs). The proposed methodology leverages unsupervised learning techniques to detect anomalies in Key Performance Indicators (KPIs), combined with explainability methods to interpret the results and highlight the most relevant contributing factors. As this is ongoing work, particular emphasis is placed on the anomaly detection component, while the root cause analysis capabilities are still being refined.

The main contributions of this work are:

- A scalable framework for anomaly detection in RAN performance data using unsupervised machine learning techniques;
- The integration of explainable AI methods to support root cause analysis, currently under development;
- An experimental evaluation demonstrating the potential of the proposed approach in real-world scenarios.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 introduces the main concepts. Section 4 presents the proposed methodology. Section 5 discusses the results, and Sect. 6 concludes the paper.

2 Related Work

Some work has been done regarding automated processes for root cause analysis in radio access networks.

Cílino et al. [1] proposed a method to identify the causes of low downlink throughput by clustering Shapley Additive Explanations (SHAP) explanations obtained from a boosting classifier, and then labelling the resulting clusters with the corresponding root causes. This allowed for classifications of the causes on the whole dataset, however not in new observations.

In [4], Gómez-Andrades et al. developed an automatic diagnosis system for Long Time Evolution (LTE) networks using a Self-Organizing Map (SOM). After training the SOM, neurons were grouped through hierarchical clustering and new observations were classified according to the cluster of their best matching unit. The method was able to capture meaningful structures in Key Performance Indicators (KPIs) and reduce manual analysis, although its performance depends heavily on feature selection and assumptions about the statistical behavior of the metrics.

Ferreira et al. [3] analyzed the root causes of low accessibility using Principal Component Analysis (PCA) combined with feature-importance metrics. Cells with Evolved Radio Access Bearer (E-RAB) setup failures above the 90th percentile were labeled as faulty. Although feature importance could be traced back to the original variables, the percentile-based fault definition and the linearity of PCA limited the effectiveness of the approach, particularly at the individual cell level.

Recent work has explored Graph Neural Network (GNN) for root cause analysis in RAN environments. The Simba framework proposed in [5] combines a graph learning module, a Graph Convolutional Network (GCN), and a transformer to capture both spatial and temporal dependencies in the KPI time series. Trained on simulated data, the model achieved high F1-scores while identifying different fault types such as interference and excessive power reduction. Similarly, Yen et al. [7] introduced a framework that integrates Graph Structure Learning with GNN models, including GCNs and Graph Attention Networks (GANs), to infer dependencies between cells and classify faults from multivariate KPI time-series data. Although these GNN-based approaches achieve strong performance, their high computational cost limits scalability to large-scale networks.

In this paper, a different approach is proposed. SHAP values are used due to their capacity to generate interpretable insights from models trained on unsupervised data. In addition to providing clear explanations of model predictions, they enable the systematic assessment of feature importance, thereby allowing for root cause analysis.

3 Preliminary Concepts

In the context of mobile communication networks, particularly in Radio Access Networks (RAN), KPIs represent measurements that reflect the performance, quality, and reliability of the network. These metrics are continuously collected from network elements and provide engineers with insights into overall network behavior. A few of common examples of RAN KPIs include:

- **Throughput** – the rate at which data is successfully transmitted;
- **Latency** – the delay experienced during data transmission;
- **Packet loss rate** – the percentage of data packets that fail to reach their destination;
- **Call drop rate** – the proportion of calls that are unexpectedly terminated.

This paper focuses on the KPI: **E-RAB Retainability−Percentage Lost**, which will be referred to as "retainability" for brevity. It measures the percentage of established E-RABs that are abnormally released, therefore lower values indicate better performance.

3.1 Anomalies

Traditionally, MNOs use simple heuristics for detecting anomalies in KPIs where data points are flagged as anomalous whenever they exceed some predefined threshold. In the case of retainability, a common threshold is 2%, however, this criterion is very far from encompassing the type of faults that are actually relevant for maintaining a well-performing network. In reality, cells have distinct temporal patterns that can't be accounted for by simple heuristics. Thus, any anomaly detection method for this application should have into consideration the seasonal patterns typical to each cell. In general, cells tend to have much lower retainability at night, therefore a value that would be considered normal at midday may be anomalous if it occurs at midnight.

Another crucial aspect to consider is that cells may degrade momentarily, producing isolated anomalies in time before returning to normal operation, and such cases do not require intervention. The focus should therefore be on degradations that persist over time, since these are the ones that require corrective action. An anomaly score function S is a function from the set of data points $\mathcal{D}$ to the real line $\mathbb{R}$ that quantifies how anomalous each data point is.

Definition 1 (Sporadic Anomaly). *For a time series* $\{x_i\}_{i=1}^n$*, let* $S : \mathcal{D} = \{x_1, \ldots, x_n\} \to \mathbb{R}$ *be an anomaly score function and let* $T \in \mathbb{R}$ *be a given threshold. A data point* $x_i \in \mathcal{D}$ *is said to be a sporadic anomaly if* $S(x_i) \geq T$*.*

Definition 2 (Persistent Anomaly). *Given a time series* $\{x_i\}_{i=1}^n$ *and a window size* $k \in \mathbb{N}$*, any data point* x_i *such that* $x_{i-k+1}, x_{i-k+2}, \ldots, x_i$ *are sporadic anomalies is said to be a persistent anomaly.*

The only cells that need intervention are those with persistent anomalies, for the reasons specified above.

3.2 Shapley Additive Explanations

Shapley Additive Explanations (SHAP) is a framework for interpreting ML models that unifies several methods from the family of additive feature attribution models. It is particularly useful for explaining predictions of complex or black-box models. SHAP provides local explanations, meaning it attributes a contribution value to each feature for a specific prediction rather than only giving a global measure of feature importance. This makes it possible to understand how individual features influence the model's decision for a given observation. The explanation model in SHAP is defined as:

$$g(z') = \phi_0 + \sum_{i=1}^{M} \phi_i z'_i, \tag{1}$$

where $z' \in \{0,1\}^M$ is a simplified binary representation of the input features, indicating the presence ($z' = 1$) or absence ($z' = 0$) of each feature. The term ϕ_0 represents the expected model output when no features are present, serving as a baseline. Each $\phi_i \in \mathbb{R}$ quantifies the contribution of feature i to the specific prediction under consideration, capturing how much that feature increases or decreases the model's output relative to the baseline. By summing these contributions, $g(z')$ approximates the original model's prediction for the given observation, allowing a clear interpretation of individual feature effects.

In the calculations of SHAP values, TreeExplainer [6] was used, which is optimized for tree-based models and offers significantly faster computation compared to the model-agnostic KernelSHAP.

4 Methodology

The dataset used includes real world data provided by Multivision, and comprises 11407 cells from a 4G network. It spans over a period of 21 d, and has granularity of 15 min. Each row in the dataset corresponds to a measurement from a specific cell at a given timestamp. Besides the cell identifier and timestamp columns, the dataset has one column for the retainability and 95 others for performance management counters, which were selected due to their potential to indicate root causes for anomalous retainability values.

The anomaly detection component described in Subsect. 4.1 operates on the time series of a single cell at a time. Prior to analysis, the time series is preprocessed by applying linear imputation to handle missing values. In contrast, the RCA model presented in Subsect. 4.2 takes as input both the explanatory performance counters and the retainability values. Unlike the AD model, no imputation is performed; instead, any rows containing missing values are removed.

4.1 Anomaly Detection

Seasonal-Trend Decomposition using LOESS (STL) [2] is a widely used statistical technique for decomposing a univariate time series into three components: trend, seasonality, and residuals (2). The trend component captures long-term changes in the series, the seasonal component captures periodic fluctuations, and the residual component contains the remaining variability not explained by the trend or seasonality. In the context of anomaly detection, the residual component is of particular interest, since it represents deviations from expected behavior after removing trend and seasonal components. STL was applied to decompose the retainability time series $\{y_i\}_{i=1}^n$ into trend t_i, seasonality s_i, and residuals r_i, such that

$$y_i = t_i + s_i + r_i. \tag{2}$$

Iterative robustness is incorporated by introducing weights that depend on the residuals and are updated across iterations. At each step, the decomposition is obtained by solving a weighted smoothing problem, where observations with large residuals are assigned smaller weights. This progressively reduces the

influence of anomalous points on the estimated trend and seasonal components, concentrating their effect in the residual term. This produces a decomposition with a clearer trend and seasonality and larger residuals for anomalous data points, facilitating their detection. Two anomaly scores were considered. The first score is the absolute value of the residual,

$$S(y_i) = |r_i|, \tag{3}$$

and the second score is the absolute value of the normalized residual,

$$\widetilde{S}(y_i) = \left| \frac{r_i - \bar{r}}{\widehat{\mathrm{Var}}(r)} \right|. \tag{4}$$

The score function $\widetilde{S}$ measures how anomalous an observation is relative to the variability of its own time series, allowing anomalies to be detected within each individual cell. However, relying solely on this score can lead to many non-relevant points being flagged as anomalies. For example, a cell with a very consistent daily cycle may be flagged as anomalous even if it deviates by a very small amount. Although this technically constitutes an anomaly relative to the cell's own behavior, such a small deviation is typically not meaningful for MNOs. For this reason, it was also necessary to consider the score S.

The data points of interest are those that are persistent anomalies according to both S and $\widetilde{S}$ relative to some thresholds $T_a \geq 0$ and $T_r \geq 0$, respectively. These thresholds are parameters of the procedure and should be chosen according to the desired sensitivity.

4.2 Root Cause Analysis

Cilínio et al. [1] categorize as anomalous any data point with throughput greater than 7 Mb/s, framing the task as a binary classification problem. However, this type of labeling is overly coarse, as it discards valuable information contained in the original KPI values. Instead of training a binary model to predict whether a data point is anomalous or not, we opt for a regression model. This approach preserves more information and provides greater flexibility in the choice of anomaly detection techniques. Therefore, a regression was performed on the raw KPI value and SHAP was applied only to those points that could be reliably predicted by the model.

First, the dataset was split into training, validation, and test sets. The split was performed temporally: the first 70% of the dataset was used for training, the next 15% for validation, and the final 15% for testing. The validations set was used for early stopping. Three gradient boosting models were trained to predict retainability: LightGBM, CatBoost, and XGBoost. Then the model achieving the lowest Root Mean Squared Error (RMSE) (5) was selected as the final predictor, however optimal predictive performance is not a strict requirement, as only the observations that are predicted with sufficient accuracy are subsequently

explained.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2} \tag{5}$$

The reliability of SHAP explanations depends on the predictive accuracy of the underlying model: the smaller the prediction error for a data point, the more accurately the model captures its behavior and, consequently, the more informative the corresponding SHAP explanations become. For this reason, the explanation phase was restricted to observations that the model predicted with sufficient accuracy. To enforce this constraint, an error threshold T_e was defined, and a data point was considered well explained whenever

$$|y_i - \widehat{y}_i| \leq T_e, \tag{6}$$

where $\widehat{y}_i$ denotes the model's prediction for the i-th observation. The value of T_e was chosen based on practical considerations, reflecting both network operational constraints and the relevant business objectives.

The SHAP values were then processed using Uniform Manifold Approximation and Projection (UMAP), a dimensionality reduction technique that preserves the global and local structure of the data while projecting it into a lower-dimensional space. After dimensionality reduction, the data was clustered using Density-Based Spatial Clustering of Applications with Noise (DBSCAN), enabling the identification of groups of observations with similar SHAP value patterns. Each resulting cluster was subsequently labeled according to the most likely underlying cause of the anomaly.

Only cases with persistent anomalies are of interest to Mobile Network Operators (MNOs), therefore in order to determine the root cause of each cluster, an analysis was conducted on only the persistent anomalies within each cluster using the most broad criteria which consists in having three consecutive sporadic anomalies. This allows for any criteria to be used, because any cell with persistent anomalies needs to have at least three consecutive sporadic anomalies.

5 Results

All experiments were conducted on a standard computing environment equipped with an Intel Core i5-2520M processor and 16 GB of RAM, without the use of specialized hardware such as GPUs.

STL was used to decompose the time series, with the period parameter set to 96, corresponding to the number of data points in a day given that the granularity of the time series is 15 min. The threshold T_r was set to 2 in order to flag as anomalous any data points that deviate from the seasonal pattern more than 2 standard deviations, and T_a was set to 0.5 so that only absolute deviations greater than that are considered. This resulted in effective anomaly detection, as illustrated by the examples in Fig. 1. Specifically, the cell with ID "UNIBNCL" exhibits only sporadic anomalies, whereas cell "SHOPNC2L" shows multiple

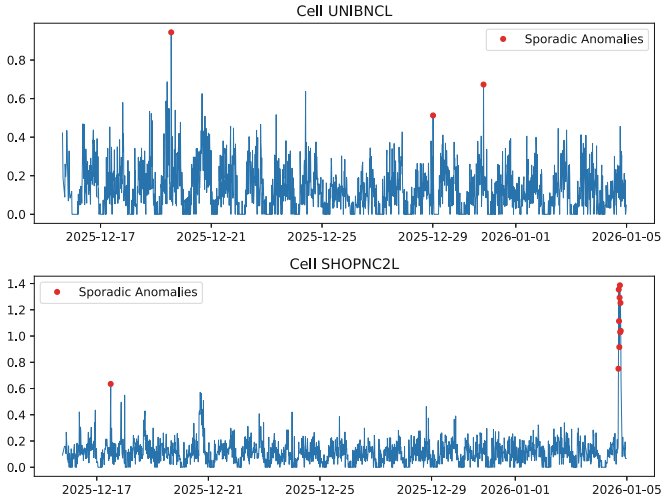


Fig. 1. Retainability time series for cells UNIBNCL (top) and SHOPNC2L (bottom) with flagged sporadic anomalies highlighted in red. (Color figure online)

sporadic anomalies toward the end of its time series, which could be classified as persistent depending on the chosen window size.

In the RCA phase, among the trained models, LightGBM achieved the lowest RMSE of 0.318, while XGBoost obtained 0.342 and CatBoost 0.357. Hence, LightGBM was selected as the final predictor.

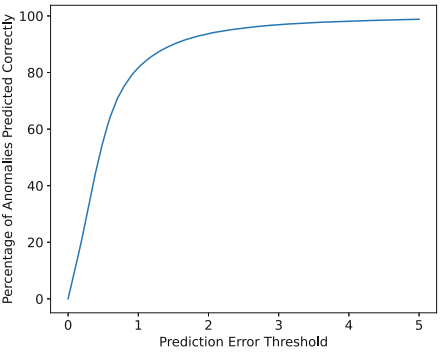


Fig. 2. Percentage of well-predicted anomalies plotted against prediction error threshold T_e.

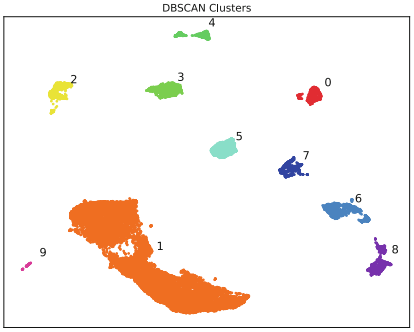


Fig. 3. Clusters of the embedded SHAP values.

An important aspect in evaluating the effectiveness of the model is the proportion of anomalies that can be explained. To illustrate how the explanation

coverage depends on the chosen threshold, Fig. 2 shows the percentage of anomalies that are well-predicted by the model for different values of the prediction error threshold. As T_e increases the percentage of well-predicted observations rises extremely quickly, meaning this threshold can be quite conservative and still capture a broad range of anomalous cases. The error threshold T_e was set to 1, under which 81% of the sporadic anomalies were well-predicted by LighGBM and could therefore be explained. After computing the SHAP values for all well-predicted anomalies, UMAP was applied, and the resulting two-dimensional representations were then clustered using DBSCAN with the parameters present in Table 1. This resulted in the clustering illustrated in Fig. 3, where 10 clearly distinct clusters are visible.

Table 1. Parameters used for UMAP and DBSCAN

Parameter	Description	Value
UMAP		
`n_neighbors`	Number of neighboring points used in local approximations of the manifold structure.	15
`min_dist`	Minimum distance between embedded points in the low-dimensional space.	0
`n_components`	Number of dimensions of the embedding space.	2
`metric`	Distance metric.	euclidean
DBSCAN		
`eps`	Maximum distance between two samples for them to be considered as neighbors.	0.8
`min_samples`	Minimum number of samples required to form a dense region (core point).	15
`metric`	Distance metric.	euclidean

Cases exhibiting three consecutive anomalies were analyzed to identify their root causes. Clusters containing persistent anomalies were clusters 1, 3, 6, 7, and 9, therefore the analysis focused exclusively on these clusters. Three distinct groups were found. One of these groups, Group A, maps directly to cluster 7, meanwhile the other two groups, Groups B and C, are spread across clusters 1, 3, 6, and 9. The root cause associated with Group A is an automated cell downtime event, consistent with a hardware or software fault activating the eNodeB supervision mechanisms within the 15-minute window. This is shown by a strong elevation in automated cell downtime, a simultaneous drop in coverage sampling, and a sharp increase in abnormal User Equipment (UE) context releases triggered by the downtime condition. Since this type of fault is indicative of a hardware issue, the radio engineer should redirect it to the field operations team for on-site investigation and corrective actions. Group B is characterised by Physical Random Access Channel (PRACH) and Contention-Based Random

Access (CBRA) congestion. The affected cells exhibit chronically low CBRA success rates well below the population median, while Contention-Free Random Access (CFRA) success rates remain near 100%. This contrast rules out a radio unit fault (where contention-free random access would also degrade under hardware or coverage failure) and points instead to saturation of the shared random access channel, most likely caused by high active UE density or a suboptimal Random Access Channel (RACH) configuration. In this context, the radio engineer should first verify the RACH configuration parameters, and if these are correctly set, the next step is to investigate the presence of temporary high UE concentration events (e.g., concerts, sports events) and apply appropriate optimization actions. Group C presents isolated retainability spikes against otherwise healthy KPI profiles, with no structural failure signature in adjacent counters. These anomalies are attributable to short-duration interference bursts or load spikes confined to the minimum detection window of 45 min and are not expected to persist under a more stringent threshold, making them candidates for exclusion from further investigation. In such cases, the radio engineer should keep the affected cells under observation to ensure that the degradation does not recur or evolve into a persistent issue.

The methodology successfully distinguished Group A, corresponding to cluster 7, from all other possible root causes. The SHAP values clustering isolated these events unambiguously because the model consistently attributed the elevated retainability prediction to downtime-related features, confirmed by the underlying counter evidence. While Groups B and C required indirect characterisation through cell-level counter analysis independently of the clustering result, the SHAP-based clusters having established only that these events were explained differently from Group A, the contrast between chronically depressed CBRA success rates and otherwise healthy structural indicators provided a principled basis for separating structural channel congestion from transient load events. The obtained execution times demonstrate the practical feasibility of the proposed methodology for real-world applications. Specifically, the anomaly detection stage required approximately 40 min, while the training phase of the root cause analysis model took around 19 min. Once trained, the RCA model exhibited a very low inference latency, requiring only 0.08 s to analyze a single data instance. These results indicate that, despite moderate offline training costs, the proposed framework is well-suited for near real-time root cause analysis in operational Radio Access Network environments, where rapid decision-making is critical.

6 Conclusion

In this paper, we presented a Machine Learning-based framework for anomaly detection and root cause analysis in Radio Access Networks. By combining unsupervised anomaly detection techniques with explainability methods, the proposed approach enables the identification of abnormal network behavior and provides meaningful insights into its underlying causes.

The results demonstrate that the framework is capable of effectively detecting anomalies in network performance data and supporting engineers in diagnosing issues more efficiently. The use of interpretability tools, such as SHAP, proved particularly valuable in understanding model decisions and highlighting the most relevant Key Performance Indicators contributing to anomalies.

Overall, this work shows that integrating anomaly detection with explainable AI can significantly enhance troubleshooting processes in modern communication networks, reducing manual effort and improving operational performance.

Future work may include the integration of real-time monitoring capabilities, the exploration of advanced deep learning models, and the extension of the approach to other domains within telecommunications networks.

Acknowledgments. This work is funded by national funds through the 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). This work was supported by the European Union via the European Regional Development Fund (ERDF) under the PORTUGAL 2030 and COMPETE 2030 Programme [COMPETE2030-FEDER-01176700].

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

References

1. Cilínio, M., Duarte, D., Vieira, P., Pereira, M., Mata, L.: Explainable fault analysis in mobile networks: a SHAP-based supervised clustering approach (2023). https://doi.org/10.1109/ICSPCS58109.2023.10261152
2. Cleveland, R.B., McRae, J.E., Terpenning, I.: STL: a seasonal-trend decomposition procedure based on loess. J. Off. Stat. **6**(1), 3–73 (1990)
3. Ferreira, D., et al.: Root Cause Analysis of Reduced Accessibility in 4g Networks. IEEE Press (2020)
4. Gómez-Andrades, A., Muñoz, P., Serrano, I., Barco, R.: Automatic root cause analysis for LTE networks based on unsupervised techniques. IEEE Trans. Veh. Technol. **65**(4), 2369–2386 (2016). https://doi.org/10.1109/TVT.2015.2431742
5. Hasan, A., Boeira, C., Papry, K., Ju, Y., Zhu, Z., Haque, I.: Root cause analysis of anomalies in 5g ran using graph neural network and transformer (2024)
6. Lundberg, S.M., et al.: From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intell. **2**(1), 2522–5839 (2020)
7. Yen, C., et al.: Graph neural network based root cause analysis using multivariate time-series KPIs for wireless networks. In: NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium. IEEE Press (2022). https://doi.org/10.1109/NOMS54207.2022.9789858

THEMATIC2GREEN: Strengthening Innovation Capacity for a Place-Based Green Transition in Fragile ADRION Territories

Michele Greco, Francesco Scorza, Biagio Perretti, Francesco Paolo Del Giudice, Alfonso Annunziata, and Beniamino Murgante(✉)

University of Basilicata, Viale dell'ateneo Lucano, 85100 Potenza, Italy
beniamino.murgante@unibas.it

Abstract. Coastal, mountainous, and historically significant territories in the ADRION macro-region face increasing pressure from environmental degradation, climate change, and uneven patterns of economic development. In these areas, small and medium-sized enterprises, public institutions, and civil society organizations must reconcile environmental sustainability with social sustainability while preserving employment and local income opportunities. In response, THEMATIC2GREEN aims to strengthen innovation capacity for green development based on cultural and natural resources. The project adopts a place-sensitive approach structured around three interdependent lines of action: environmental quality management and certification of economic activities, climate change adaptation, and renewable energy production and consumption, including energy communities. Targeting fragile coastal and mountainous regions, the project combines pilot actions, joint solutions, a joint strategy for renewable energy communities, and a permanent cooperation network. Through a quadruple helix model, THEMATIC2GREEN seeks to generate practical, transferable, and territorially adapted solutions for a balanced green transition in the ADRION area.

Keywords: green transition · innovation capacity · fragile territories · sustainable development

1 Introduction

In the coming years, small and medium-sized enterprises (SMEs), public institutions, and non-governmental organizations (NGOs) located in coastal areas, mountain territories, and historic towns of lagging regions will face a common set of major challenges [1, 2]. These communities must reconcile environmental sustainability with social sustainability, while preserving local employment and income opportunities [3, 4].

The first issue concerns the prevention and reduction of degradation of natural resources and cultural heritage caused by economic activities that remain essential for local development. In these territories, agriculture, tourism, and other place-based services provide livelihoods but can also generate environmental pressure [3, 5, 6]. The key

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 270–280, 2027.
https://doi.org/10.1007/978-3-032-30539-8_20

issue is therefore to improve the environmental performance of local economic systems without weakening their social and economic functions. A second challenge is the transition away from fossil fuels while improving the efficiency and sustainability of renewable energy production and distribution [7–10]. This issue is particularly complex in marginal areas characterized by low population density and high environmental and cultural value. In such territories, renewable energy systems often involve high production and distribution costs. In addition, renewable sources may also generate significant environmental impacts, including visual impacts on landscapes from wind turbines and solar installations, pressure on air quality and agricultural land from biomass plants, and landscape fragmentation associated with energy distribution networks. As a result, the energy transition in these areas requires solutions that are both technically efficient and territorially sensitive. A third challenge is adaptation to climate change and mitigation of its effects on local economic and social activities. Climate change is already affecting, or is expected to affect in the short and medium term, sectors such as cultural services, cultural and nature-based tourism, non-food agricultural production, and other consumer-oriented services. SMEs, public and private service providers, and local communities will need to adjust production processes, services, and product innovation strategies in order to reduce vulnerability and strengthen resilience [11–20]. At the same time, this context creates important opportunities. These territories are increasingly included in European and macro-regional policy frameworks that support a sustainable transition, notably the European Green Deal and the EUSAIR strategy. In this perspective, the development of environmentally high-quality businesses, innovative systems for renewable energy production and consumption, and stronger local capacities to mitigate and adapt to climate change can generate both local and global benefits, including new employment opportunities. Against this background, the overall objective of THEMATIC2GREEN is to strengthen innovation capacity for green development based on cultural and natural resources, while addressing emerging environmental and social challenges. The project is designed to support territorial development in the ADRION macro-region through coordinated action in three interdependent thematic areas. First, the project aims to enhance entrepreneurial capacity to improve the environmental sustainability of local economic systems. This includes promoting environmental quality, innovation, and certification processes for economic activities rooted in local resources. Second, it aims to reinforce the capacity of public and private actors to identify climate-related threats, assess local impacts, and define adaptation strategies for preserving both economic development and the social sustainability of the green transition. Third, the project promotes the development of renewable energy production, distribution, and consumption, with particular attention to the creation of locally adapted and socially acceptable solutions, including a joint strategy for renewable energy communities. The main expected outputs include pilot actions, joint solutions, a joint strategy for renewable energy community development, and a permanent cooperation network able to continue beyond the project lifetime. The project focuses on coastal and mountainous areas in the ADRION macro-region, which share rich natural resources, diverse cultural heritage, small rural populations, and economies centered on agriculture, tourism, and micro-enterprises. These regions are also highly vulnerable to climate change impacts such as sea-level rise, erosion, and extreme weather. The partnership of THEMATIC2GREEN higher education institutions,

business support agencies, local administrations, and development agencies across the ADRION area. Based on a quadruple helix model, the project integrates academia, public institutions, businesses, and civil society to support the green transition through three lines of action: environmental quality in economic activities, renewable energy systems, and climate change adaptation. Its integrated and place-sensitive approach aims to generate practical and transferable solutions for fragile territories.

2 Organization of the THEMATIC2GREEN Project

2.1 Methodological Approach

THEMATIC2GREEN adopts an integrated territorial approach to foster innovation capacity for green development in marginal coastal and mountainous areas characterized by rich cultural and natural resources. The method is designed to support public institutions and micro-scale public and private businesses, which often operate under limited human resources and therefore face difficulties in addressing green-transition objectives through isolated sectoral interventions. The project addresses three interrelated dimensions in parallel. The first concerns the improvement of environmental quality management in economic activities, particularly those based on cultural and natural heritage. The second focuses on innovative models for renewable energy production, distribution, and consumption, with specific attention to locally adapted solutions such as energy communities. The third aims to strengthen the capacity of public and private actors to identify climate-related risks, assess local impacts, and define adaptation strategies able to preserve economic development and social sustainability.

The methodological assumption is that these dimensions are mutually dependent and should be implemented through an integrated framework. This design is intended to maximize systemic impact, promote synergies among stakeholders, reduce fragmentation in project coordination, and avoid excessive pressure on the limited administrative and organizational capacities of marginal territories.

Operationally, the project combines pilot actions, joint solutions, strategic planning, and network-building activities. In each thematic area, partners co-develop and test solutions tailored to local needs, while also contributing to a joint strategy for renewable energy communities. These activities are supported by the creation of a permanent cooperation network aimed at ensuring continuity beyond the project duration.

The methodological framework also capitalizes on the experience developed in the previous THEMATIC project under the ADRION 2014–2020 programme. THEMATIC2GREEN builds on that experience by reusing knowledge on territorial needs, local potential, and stakeholder networks, thereby strengthening complementarities among partners and consolidating already established cooperation mechanisms.

The method also leverages partners' prior expertise. In environmental management and certification, partners share capacity-building experience and help transfer good practices. For renewable energy, operational knowledge is exchanged on innovative local models like energy communities. Climate adaptation solutions are developed using insights from earlier national and European projects involving multiple partners.

Overall, THEMATIC2GREEN is based on integration, capitalization of previous experience, and collaborative experimentation, with the aim of generating practical

and transferable solutions for a balanced and effective green transition in vulnerable territories.

2.2 Work Plan and Work Packages

The project work plan is structured into three thematic Work Packages (WPs), corresponding to the main dimensions of the green transition addressed by THEMATIC2GREEN: environmental quality of economic activities, climate change adaptation, and renewable energy systems. Across all WPs, the project follows a shared methodological sequence: territorial and needs-based analysis, co-design and implementation of pilot actions, and dissemination and transfer of results. This common structure ensures coherence and supports the progressive strengthening of institutional and operational capacity.

WP1, Promotion of environmental quality certification of public and private services, focuses on improving the capacity of public and private actors in coastal and mountainous territories whose economies rely on cultural and natural resources. Its target groups are SMEs, public institutions, and NGOs involved in cultural services and heritage-related activities. The first action develops a structured analysis of territorial needs and potentials through a transnational working group, surveys, and the identification of support options, leading to a shared white paper. The second action implements a pilot aimed at creating a technical assistance service for environmental management and certification, including a feasibility study, a support platform with models and diagnostic procedures, direct assistance to final beneficiaries, staff training, and a transferable joint solution. The third action addresses communication and awareness raising through workshops, conferences, social media, and video dissemination. The main output is a support platform for environmental management and certification, accompanied by a joint solution for reuse in partner regions.

WP2, Pilot actions for creation of tools for climate change adaptation of vulnerable economic systems, aims to strengthen the ability of public and private actors to assess and respond to climate-related impacts in vulnerable coastal and inland areas. It addresses tourism, cultural services, archaeological and natural heritage, and rural communities. The first action consists of a territorial analysis of expected climate impacts on local economic systems, based on the review of available models, the selection of indicators, and stakeholder surveys. The second action develops a pilot based on the joint design of local-scale forecast models, supported by GIS tools and adapted to partner territories. These models are conceived as decision-support tools for investment planning, territorial management, and innovation. The third action disseminates adaptation options through workshops, video production, and social media campaigns. The main outputs are a forecast platform for local climate scenarios and a joint solution for cooperative adaptation strategies (Fig. 1).

WP3, Pilot actions and joint solutions for energy efficiency in coastal and rural areas, addresses renewable energy development and energy efficiency in low-density, fragile, and seasonally variable territories. The first action combines territorial analysis and stakeholder consultation for the preparation of a Joint Strategy for the promotion of energy communities, while also supporting the formalization of a permanent cooperation network. The second action consists of joint pilot activities on micro-installations

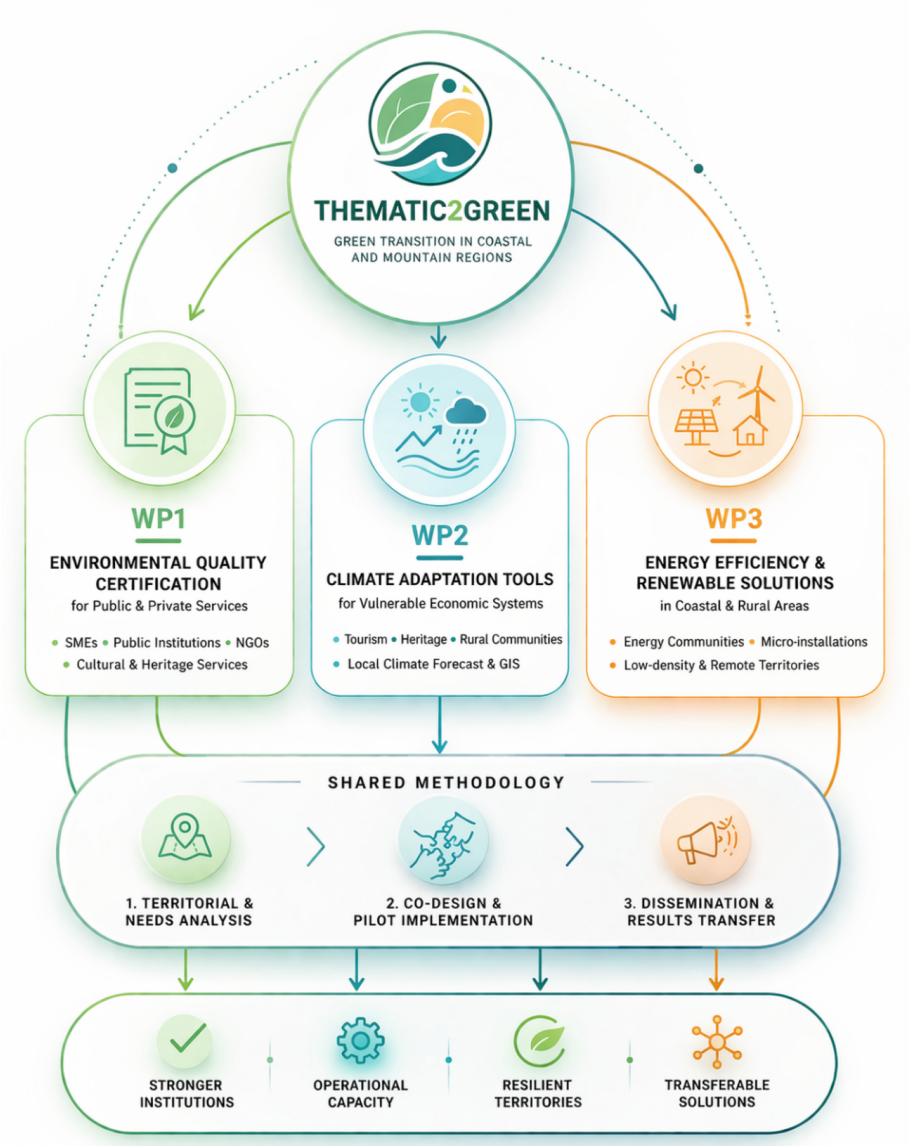


Fig. 1. Project Structure and Articulation into Work Packages

and innovative renewable energy systems in coastal, rural, and remote contexts. The third action promotes awareness raising and mainstreaming through communication campaigns, newsletters, international fairs, and European sustainability events.

Overall, the work plan presents a coherent and methodologically robust design, in which each thematic line moves from diagnosis to experimentation and from experimentation to transferability.

2.3 Structure of the Partnership

The partnership is structured according to the triple helix model, combining research and innovation institutions, public authorities responsible for policy design and implementation, and organizations specialized in business support and promotion. This composition is intended to ensure both scientific rigor and practical relevance in the development of actions for the green transition in the ADRION macro-region. Beyond project outputs, the consortium is designed to generate a lasting impact at both regional and transnational scales by consolidating a network able to support coordinated green development across diverse territorial contexts. A major strength of the partnership lies in the capitalization of experience gained through the previous THEMATIC project (ADRION 2014–2020), in which most partners worked together over a three-year period. That cooperation continued after the formal end of the project through a permanent stakeholder network established by a memorandum of understanding signed in 2021. This network played an important role in the development of the THEMATIC2GREEN proposal and provides a strong institutional basis for the new project. The partnership combines four complementary dimensions: scientific and technical expertise, geographical representativeness, policy-making capacity, and business development experience. Scientific and technical expertise is mainly provided by the academic partner SI-UNIBAS, together with local development agencies and business support organizations. Geographical coverage is ensured by the presence of partners from four ERDF countries and three IPA countries, representing a wide range of environmental, structural, and economic conditions within the Adriatic-Ionian area. Policy-making capacity is provided by local administrations and regional agencies, while business development experience is ensured by national and regional support bodies.

All partners contribute to project activities according to their institutional profile and field of competence. In the work package on environmental quality development and certification, ENTGRE, RDA BKP, and ZRDASER play a leading role, especially in the design and implementation of pilot actions. In the climate adaptation work package, key functions are performed by RUTGRE, SI-UNIBAS, MUNPREK, ZRDASER, BSC KRANJ, REDAH, and ENTGRE, with particular relevance for the development of models and operational tools for vulnerable mountain areas. In the renewable energy work package, leadership is mainly provided by MUNPREK, MUNMED, and BSC KRANJ, while the Municipality of Castelmezzano contributes an established good practice on energy communities in environmentally sensitive territories. All partners also contribute to the Joint Strategy through territorial analysis and feasibility assessment. Overall, the partnership reflects a balanced combination of expertise, territorial coverage, and institutional continuity. This configuration strengthens implementation capacity and enhances the consortium's ability to generate coordinated and durable effects on the green transition in the ADRION macro-region.

3 Objectives and Expected Results

3.1 Common Challenges, Strategic Alignment and European Policy Synergies

The ADRION countries present common sustainability issues despite substantial socio-economic differences. Coastal and inland rural areas face increasing pressure from resource exploitation and climate change, while the economic potential of cultural and natural assets remains unevenly developed across the macro-region. In this context, THEMATIC2GREEN aims to reinforce innovation capacity for a balanced green transition by fostering transnational cooperation on environmental quality management, renewable energy, and climate change adaptation in coastal and rural economic systems. Within this perspective, THEMATIC2GREEN is explicitly aligned with the European Green Deal, the EU Strategy for the Adriatic and Ionian Region (EUSAIR), and relevant national Smart Specialisation Strategies (S3). With respect to the Green Deal, the project contributes to decoupling economic growth from resource use and to a just transition that "leaves no person and no place behind" through actions on energy efficiency, renewable energy systems, energy communities, and environmental quality certification in lagging coastal and rural areas. In relation to EUSAIR, it supports regional cooperation, Western Balkans integration, economic development, and the valorisation of the macro-region's cultural heritage. The project is also designed to reinforce national and regional S3 priorities in partner countries, especially in the fields of sustainable energy, environment, circular economy, tourism, and innovation for green development.

3.2 Expected Results and Target Beneficiaries

The expected results of THEMATIC2GREEN derive from its comparative and transnational approach. By examining structurally similar territories across the ADRION area, the project aims to produce a more accurate understanding of common needs, constraints, and development opportunities. This approach is expected to improve diagnostic capacity and to support the joint development of robust and transferable solutions. In particular, the exchange of good practices among territories with comparable economic structures and environmental conditions should accelerate the capitalization of innovative methodologies in two key fields: environmental quality management for micro-businesses and small-scale renewable energy systems. At the same time, the joint development of models and methodologies for assessing climate change impacts is expected to strengthen the robustness and operational effectiveness of local adaptation strategies. A central expected outcome is the reinforcement of planning and implementation capacity across public authorities. Local authorities in coastal and rural areas are expected to acquire tools for investment planning in climate adaptation, renewable energy, and environmentally sustainable development. Regional authorities should benefit from improved instruments for territorial planning and investment prioritization, supported by the involvement of regional bodies and development agencies and by the capitalization of networks and results from the previous THEMATIC project. The project is also expected to strengthen the engagement of national public authorities, particularly in IPA countries, through closer technical and institutional interaction with European partners.

Sectoral agencies are also expected to play a strategic role as intermediaries between territorial needs and the joint solutions developed at transnational level. Their contribution is essential for translating project outputs into operational support mechanisms and for adapting common methodologies to local governance and market conditions. Higher education and research organizations, in turn, are expected to contribute to the identification and validation of innovative solutions, particularly in relation to renewable energy production and distribution systems suited to coastal areas with seasonal demand patterns and inland territories characterized by low population density and vulnerable ecosystems. Their involvement reinforces the scientific basis of the project and supports methodological rigor in the design of technical solutions.

SMEs represent another key target group. In rural, coastal, and mountain service sectors, the project is expected to increase awareness and practical capacity in environmental quality management, while improving access to environmental certification procedures for activities related to cultural services and the management of cultural and historical resources. The general public is also expected to benefit from greater awareness of energy communities and of their social, economic, and environmental advantages, which is a necessary condition for their wider diffusion.

Overall, the project aims to support a multi-level and transnational learning process for a more coherent and inclusive green transition across the ADRION macro-region.

4 Preliminary Results

Building on the objectives and expected results outlined in Sect. 3, the Metaponto case study provides an initial empirical validation of the project's analytical framework by showing how compound coastal hazards can progressively reconfigure territorial risk in fragile Mediterranean systems. The assessment, based on a multi-scenario approach integrating relative sea level rise (RSLR), vertical land movements (VLM), and storm surge dynamics, identifies a distinctly non-linear evolution of impacts, moving from selective exposure in the short term to accelerated damage in the medium term and systemic risk amplification in the long term.

In the 2030 horizon, impacts remain relatively contained and spatially differentiated. The most exposed component is the drainage network, with interference values between approximately 18% and 25% of its total extent, while protected areas and the road network show lower but still significant levels of impact. By contrast, buildings and accommodation facilities remain only marginally affected in most scenarios. This pattern indicates that early climate-related pressures primarily affect the environmental and functional backbone of the territorial system, particularly regulation and drainage infrastructures, before significantly compromising the built environment and economic assets.

A marked discontinuity emerges by 2050. Under intermediate and high-impact scenarios, damage and interference indicators increase sharply across several territorial components, especially accommodation facilities, buildings, and protected areas, with values entering the 30–40% range. This acceleration suggests progressive saturation of the system's adaptive capacity and reflects the cumulative interaction of sea-level rise, subsidence, and more intense extreme events. In this perspective, 2050 can be interpreted

as the first critical tipping point, beyond which the territorial system loses a substantial part of its capacity to absorb external perturbations without structural and functional reconfiguration.

By 2100, the analysis delineates a clearly systemic risk scenario. Under high-forcing conditions (RCP8.5 combined with VLM), interference reaches about 37.3% for the drainage network, 44.4% for protected areas, 26% for accommodation facilities, 11.3% for buildings, and 12.1% for the road network, while irrigated areas remain comparatively less exposed, although still affected by increasing trends. The inclusion of storm surge dynamics produces a further step change in impact magnitude: interference exceeds 60% for accommodation facilities and approaches 70% for protected areas even under moderate return-time conditions, with peaks close to 80% under the most severe compound scenario (RCP8.5 + VLM combined with a 100-year storm surge event). These results indicate that extreme events do not act as simple additive stressors, but as discontinuity factors capable of triggering rapid transitions toward high-impact states.

Overall, the preliminary results reveal a clear hierarchy of vulnerability. Natural systems and environmental infrastructures act as primary receptors of climate stress, followed by a progressive propagation of impacts toward the built environment and, subsequently, toward economic activities, particularly coastal tourism. This pattern is reinforced by multi-risk interactions, as coastal erosion, flooding, groundwater salinization, and water stress operate in a mutually amplifying way. The Metaponto case therefore illustrates the transition from sectoral vulnerability to systemic risk in Mediterranean coastal territories and highlights the need for timely, integrated, and cross-sectoral adaptation strategies.

5 Conclusions

THEMATIC2GREEN addresses the green transition in fragile ADRION territories through an integrated and place-based framework that combines environmental quality management, renewable energy development, and climate change adaptation. Its methodological value lies in the parallel treatment of these interdependent dimensions through pilot actions, joint solutions, strategic planning, and transnational cooperation. The project is further strengthened by a multi-actor partnership and by the capitalization of knowledge, networks, and operational experience generated in the previous THEMATIC initiative. The preliminary evidence from the Metaponto case study confirms the relevance of this approach. The results show that climate-related pressures in Mediterranean coastal systems evolve from selective and sector-specific impacts to systemic risk configurations driven by the interaction of sea-level rise, subsidence, and extreme events. This finding supports the need for integrated, cross-sectoral, and territorially differentiated adaptation strategies.

Overall, THEMATIC2GREEN provides a structured framework for reinforcing innovation capacity in vulnerable regions and for generating practical, transferable solutions consistent with European and macro-regional sustainability objectives. In this sense, the project contributes both to local resilience building and to broader transnational learning for a balanced green transition in the ADRION area.

Acknowledgments. This study was supported by the European Union through the Interreg IPA Adrion Program within the project "Thematic Innovation capacity toward green development - THEMATIC2GREEN" (CUP: E37F24000020007).

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

References

1. Annunziata, A., Scorza, F., Corrado, S., Murgante, B.: Unveiling intra-rural divides: investigating decline and prosperity in rural areas. The case study of southern Italy. Eur. Plan. Stud. **32**, 1478–1505 (2024). https://doi.org/10.1080/09654313.2024.2335312
2. Murgante, B., Annunziata, A., Tonini, M.: Developing a taxonomy framework for assessing human capital provision: A case study of southern Italian municipalities. Appl. Geogr. **179** (2025). https://doi.org/10.1016/j.apgeog.2025.103640
3. Macalister, F.: Preparing for the future: mitigating disasters and building resilience in the cultural heritage sector. J. Inst. Conserv. **38**, 115–129 (2015). https://doi.org/10.1080/19455224.2015.1068201
4. Bosello, F. et al.: Economy-wide impacts of climate mitigation and adaptation strategies across European regions. In: Sanderson, H., Hildén, M., Russel, D., Penha-Lopes, G., Capriolo, A. (eds.) Adapting to Climate Change in Europe. Elsevier Inc (2018). https://doi.org/10.1016/B978-0-12-849887-3.00005-8
5. Sesana, E., Gagnon, A.S., Bonazza, A., Hughes, J.J.: An integrated approach for assessing the vulnerability of world heritage sites to climate change impacts. J. Cult. Herit. **41**, 211–224 (2020). https://doi.org/10.1016/j.culher.2019.06.013
6. Carroll, P., Aarrevaara, E.: Review of potential risk factors of cultural heritage sites and initial modelling for adaptation to climate change. Geosciences (Basel). **8**, 322 (2018). https://doi.org/10.3390/geosciences8090322
7. Scorza, F.: Towards self energy-management and sustainable citizens' engagement in local energy efficiency agenda. Int. J. Agric. Environ. Inf. Syst. **7** (2016). https://doi.org/10.4018/IJAEIS.2016010103
8. Corsini, F., Certomà, C., Dyer, M., Frey, M.: Participatory energy: research, imaginaries and practices on people' contribute to energy systems in the smart city. Technol. Forecast. Soc. Change. **142**, 322–332 (2019). https://doi.org/10.1016/j.techfore.2018.07.028
9. Agostini, A., Colauzzi, M., Amaducci, S.: Innovative agrivoltaic systems to produce sustainable energy: an economic and environmental assessment. Appl. Energy. **281**, 116102 (2021). https://doi.org/10.1016/j.apenergy.2020.116102
10. Waheed, R., Chang, D., Sarwar, S., Chen, W.: Forest, agriculture, renewable energy, and CO2 emission. J. Clean. Prod. **172**, 4231–4238 (2018). https://doi.org/10.1016/j.jclepro.2017.10.287
11. Palutikof, J.P., Street, R.B., Gardiner, E.P.: Decision support platforms for climate change adaptation: an overview and introduction. Clim. Chang. **153**, 459–476 (2019). https://doi.org/10.1007/s10584-019-02445-2
12. Haugen, A., Mattsson, J.: Preparations for climate change's influences on cultural heritage. Int. J. Clim. Chang. Strateg. Manag. **3**, 386–401 (2011). https://doi.org/10.1108/17568691111175678
13. European Commission: Forging a Climate-Resilient Europe - the New EU Strategy on Adaptation to Climate Change. (2021).
14. Italian Ministry for the Environment: National Adaptation Strategy to Climate Change. (2015).

15. Annunziata, A. et al.: Climate actions handbook supporting decision making for regional adaptation strategies in Basilicata, Italy. In: Gervasi, O., Murgante, B., Garau, C., Karaca, Y., Lago, M.N.F., Scorza, F., Braga, A. (eds.) Proceedings of the 25th International Conference on Computational Science and its Applications (ICCSA 2025). Springer Nature, Switzerland AG, Istanbul, Türkiye (2026). https://doi.org/10.1007/978-3-031-97651-3_22
16. Hendawy, E. et al.: The prediction and assessment of the impacts of soil sealing on agricultural land in the North Nile Delta (Egypt) using satellite data and GIS modeling. Sustainability. **11**, 4662 (2019). https://doi.org/10.3390/su11174662
17. Molaei Qelichi, M., Murgante, B., Yousefi Feshki, M., Zarghamfard, M.: Urbanization patterns in Iran visualized through spatial auto-correlation analysis. Spat. Inf. Res. **25**(5), 627–633 (2017). https://doi.org/10.1007/s41324-017-0128-0
18. Manganelli, B., Murgante, B., Saganeiti, L.: The social cost of urban sprinkling. Sustainability. **12**, 2236 (2020). https://doi.org/10.3390/su12062236
19. Partheepan, K., Musthafa, M.M., Bhavan, T., Murgante, B.: Harmonizing ecosystem services assessments with sustainable development goals: A case study of the Batticaloa lagoon watershed, Sri Lanka. City Territ. Archit. **12**(26), 2195–2701 (2025)
20. Molaei Qelichi, M., Ghasemi, K., Tepei, E., Murgante, B.: Mapping trends in urban livability research: A comprehensive bibliometric analysis. J. Plan. Lit. **39**(1), 1–15 (2025). https://doi.org/10.1177/08854122241312181

A Unified Energy Decomposition and Delocalization Approach to Rare-Gas Excimer Bonding

Marcos Mandado(✉)

Department of Physical Chemistry, University of Vigo, Lagoas-Marcosende s/n, 36310 Vigo, Spain
mandado@uvigo.es

Abstract. The nature of the bonding in rare-gas excimers is examined through the combination of interaction energy decomposition and electron delocalization analysis. Herein, the Deformation Density-based Energy Decomposition Analysis (DD-EDA) framework is extended to the study of intermolecular interactions in electronically excited state complexes. Within this framework the total interaction energy is decomposed into electrostatic, intermolecular exchange, spin repulsion and polarization energies. This interaction energy decomposition is applied to Ne_2 and Kr_2 excimers in their lowest-lying excited states and the results are compared to those obtained for the corresponding geometrically unrelaxed ground-state complexes. Even though all the energy terms increase upon excitation, the polarization energy clearly emerges as the main stabilizing component. Visualization of the polarization density reveals a pattern that differs significantly from that exhibited by highly dispersive interactions in ground state, reflecting the dominant effect of excitonic resonance in the excimer polarization, in agreement with previous works. Analysis of the Domain-Averaged Fermi Hole Density (DAFH) and electron delocalization indices (DIs) reveals the excitonic resonance also leads to a substantial increase in the interatomic electron delocalization, with DI values that agree with a marked covalent character, in contrast with the small negative values obtained in the unrelaxed ground-state complexes. These results show that the strong binding in rare-gas excimers arises from the interplay between enhanced non-covalent interactions and excitation-induced electron delocalization (covalency).

Keywords: Excimers · Excited-State Complexes · Energy Decomposition Analysis · TDDFT · Electron Delocalization

1 Introduction

Excimers are molecular complexes that are bound exclusively in electronically excited states and dissociate upon relaxation to the ground state [1]. This unusual bonding mechanism gives rise to characteristic spectroscopic signatures, including broad and red-shifted emission bands, and plays a central role in many photophysical and photochemical processes [1–6]. Because their stability emerges only upon excitation, excimers provide a unique framework for exploring intermolecular interactions beyond the ground-state paradigm.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 281–290, 2027.
https://doi.org/10.1007/978-3-032-30539-8_21

Rare-gas excimers have been extensively employed in laser technologies [7, 8] and constitute prototypical systems for investigating the nature of excited-state bonding [9–11]. In their ground state, rare-gas atoms interact only weakly through dispersion forces, leading to very shallow potential energy surfaces [12]. Upon excitation, however, they form strongly bound dimers with significantly reduced equilibrium distances and large binding energies [13]. This dramatic change in interaction strength highlights the fundamental role of electronic excitation in modifying intermolecular forces.

Understanding the origin of this stabilization requires a detailed analysis of the different contributions to the interaction energy [11]. Traditional descriptions often emphasize the role of dispersion forces in weakly bound systems. However, in excited states, additional effects arise from electronic exchange, and charge and excitonic resonances between equivalent configurations [14, 15]. These effects can lead to significant electron delocalization between fragments, introducing a degree of covalent character into the interaction.

Previous theoretical studies have shown that helium excimer exhibit strong exchange–correlation energy and bond order, challenging the conventional view of these systems as purely van der Waals complexes and suggesting covalent nature [9, 10]. In particular, the resonance character of the excited-state wavefunction implies that the excitation may be delocalized over both atoms, giving rise to bonding mechanisms that cannot be captured by classical noncovalent interactions alone.

In this work, we present an investigation of the nature of bonding in rare-gas excimers by combining energy decomposition analysis within the DD-EDA framework [16] with complementary descriptors of intermolecular electron delocalization. This approach enables a unified interpretation of excimer stabilization in terms of both energetic contributions and electronic structure, providing new insight into the fundamental mechanisms governing excited-state intermolecular interactions.

2 Theoretical Section

2.1 DD-EDA for Excited-State Complexes

Within the DD-EDA framework, the intermolecular interaction energy is decomposed into physically meaningful contributions, namely, electrostatic (elec), intermolecular exchange (exch), spin-repulsion (rep) and polarization:

$$E_{int} = E_{elec} + E_{exch} + E_{rep} + E_{pol} \tag{1}$$

where the polarization term encompasses induction and dispersion contributions. However, a rigorous separation of polarization into induction and dispersion is only possible for weak interactions where second-order perturbation effects are sufficient and higher-order terms are negligible [16]. This is not the case in rare-gas excimers, so that we will not perform the decomposition of the polarization energy into induction and dispersion. Moreover, the sum of the intermolecular exchange and spin-repulsion contributions corresponds to the well-known Pauli repulsion energy, which can be examined separately within the DD-EDA framework [16]. It is worth noting that the term intermolecular exchange denotes the attractive component of the Pauli energy, arising—as

the name suggests—from electronic exchange between different monomers. In contrast, spin-repulsion refers to the repulsive component of the Pauli energy, originating from changes in the one-electron density induced by the Pauli exclusion principle.

The formalism is applicable to both idempotent and non-idempotent density matrices. In the case of excited states described within TDDFT, the density matrix is non-idempotent and can be diagonalized to yield natural orbitals and occupation numbers. This allows the extension of the DD-EDA scheme to excited states without modification of its fundamental structure.

A key aspect of the extension concerns the definition of the non-interacting reference system. For excimers formed by identical fragments (monomers A and B), the excited state is described by a symmetric combination of configurations in which the excitation is localized on either fragment. The corresponding electron density for the non-interacting system is then constructed as an average of ground- and excited-state (GS and ES) monomer densities, ensuring a consistent and physically meaningful reference.

$$\langle \rho \rangle = \frac{1}{2}\left(\left\langle \rho_A^{\mathrm{GS}} \right\rangle + \left\langle \rho_A^{\mathrm{ES}} \right\rangle + \left\langle \rho_B^{\mathrm{GS}} \right\rangle + \left\langle \rho_B^{\mathrm{ES}} \right\rangle\right) \tag{2}$$

The electrostatic energy is evaluated using these averaged densities, while exchange and spin-repulsion contributions are obtained through weighted orthogonalization of the monomers' natural orbitals [16]. This procedure provides a direct and internally consistent way of analyzing intermolecular interactions in excited states.

2.2 Intermolecular Electron Delocalization

To characterize the covalent component of the interaction, intermolecular electron delocalization was analyzed using both qualitative and quantitative descriptors.

Domain-Averaged Fermi Hole (DAFH) densities provide a spatial representation of electron delocalization by describing the extent to which the exchange–correlation hole associated with one fragment extends into the domain of the other. For excited states in this work, the one-electron approximation proposed by Cooper and Ponec based on natural orbitals was employed [17],

$$g_\Omega(\mathbf{r}) = \sum_{i,j} n_{ij}(\Omega)\eta_i(\mathbf{r})\eta_j(\mathbf{r}) \tag{3}$$

where η_i denote the natural orbitals and $n_{ij}(\Omega)$ is the domain occupation matrix, defined as

$$n_{ij} = \sqrt{n_i n_j} \sum_k S_{ik}(\Omega) S_{kj}(\Omega) \tag{4}$$

Here n_i are the natural occupation numbers and S represents the overlap integrals between natural orbitals evaluated within the spatial domain Ω.

Delocalization indices (DIs) were used as quantitative measures of electron sharing between fragments. In the case of rare-gas excimers, DIs were computed between the two atoms forming the dimer. As will be shown in the next section, large positive values

were found for the excited-state dimers, indicating large electron sharing and covalent character, whereas small negative values were obtained for the ground state dimers in the excited state geometries.

Within the one-electron approximation, which is also adopted here for the calculation of DIs in the excited states, delocalization indices are given by

$$\mathrm{DI}_{\Omega,\Omega'} = \sum_{i,j} \sqrt{n_i n_j} S_{ij}(\Omega) S_{ji}(\Omega') \tag{5}$$

2.3 Computational Details

The Ne_2 and Kr_2 excimers were optimized in their lowest-lying excited state (S1) using TDDFT with the M06-2X functional and the aug-cc-pVTZ basis set, including D3 dispersion correction. Basis-set effects were analyzed using the aug-cc-pVXZ series and the basis set superposition error (BSSE) evaluated using the counterpoise (CP) correction [18]. Electronic structure calculations were performed using Gaussian 16 software package [19].

Interaction energies and their components were evaluated for the excited state complexes, as well as for the geometrically unrelaxed ground state complexes (GS/S1) obtained upon vertical electronic relaxation, using the DD-EDA scheme within a locally modified version of the EDA-NCI software [20] that implements the calculation of excited state complexes. DAFH densities and DIs were computed using local algorithms from natural orbitals obtained in the electronic structure calculations. Both Mulliken and QTAIM atomic domains were employed in the calculations.

Polarization, spin-repulsion and DAFH densities were visualized and the corresponding images created using GaussView 6 [21].

3 Results and Discussion

3.1 Structure and Interaction Energies

The symmetry of the lowest-lying excited state of Ne_2 and Kr_2 corresponds to Σ_u^+. The optimized geometries are in good agreement with previous theoretical studies reporting ab initio high-level calculations, confirming the reliability of the computational approach employed in this work. Thus, the Ne–Ne and Kr-Kr equilibrium distances obtained here are 1.78 and 2.72 Å, respectively, only slightly larger than the values of 1.71 and 2.65 Å reported using the CC2 method [22]. As expected, excitation leads to significantly shorter interatomic distances compared to the ground state, reflecting the emergence of strong attractive interactions.

A critical aspect of the analysis concerns the treatment of basis-set effects. The interaction energies obtained for the aug-cc-pVXZ (with X = 2–5) basis set series with and without CP correction are represented in Fig. 1. The results show that CP correction leads to a systematic underestimation of binding energies in the excited states, yielding values that deviate significantly from the CC2 reference data (−19.60 and −20.98 kcal/mol for Ne_2 and K_2, respectively) [22]. In contrast, uncorrected interaction energies exhibit

only minor variations with basis-set size and are close to the previously reported data, indicating an effective cancellation between BSSE and intrinsic basis-set incompleteness. It must be noticed that BSE arises from the use of finite basis sets, whereas BSSE originates from evaluating the dimer in the full basis while describing the monomers with their individual basis sets. Consequently, BSSE leads to an overestimation of the IE, whereas BSE typically results in its underestimation. All of this suggests that, for rare-gas excimers, the use of CP-uncorrected interaction energies provides a more reliable description of binding, particularly when diffuse basis functions are employed.

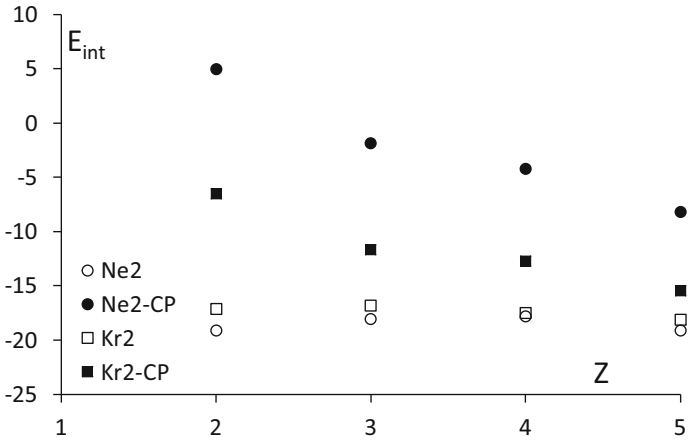


Fig. 1. IEs (in kcal/mol) obtained for the Ne_2 and Kr2 (squares) excimers using the aug-cc-pVXZ basis set series with and without CP correction at the TDDFT/M06-2X-D3 level.

3.2 Energy Decomposition Analysis

The DD-EDA results collected in Table 1 reveal a substantial increase in all interaction energy components upon excitation from GS/S1 to S1. This behavior reflects the increased overlap between the atomic densities in the excited state.

Electrostatic interactions become significantly more attractive, highlighting the importance of charge penetration effects. At the same time, the spin repulsion increases due to enhanced overlap, although this increase is partially offset by a corresponding rise in the attractive intermolecular exchange contribution. This large increase in the spin repulsion energy is qualitatively reflected on the spin repulsion densities depicted in Fig. 1 for the S1 and GS/S1 states. As anticipated, the spin-repulsion density shows a redistribution of electron density away from the intermolecular region—where orbital overlap is greatest—toward the atomic centers. This pattern becomes much more pronounced in the excited states, in line with the corresponding energy trends, which indicate stronger repulsion compared to the ground states. The effect is especially striking in Ne_2, where the depletion of electron density not only occurs between the atoms but also extends into the outer regions surrounding them, underscoring the significant interpenetration of the Ne atoms in the excimer.

Table 1. Interaction energy decomposition analysis performed within the DD-EDA framework for the Ne_2 and Kr_2 complexes in their lowest-lying excited state (S1) and in the unrelaxed ground state obtained upon vertical electronic emission (GS/S1).

	Ne_2		Kr_2	
	S1	GS/S1	S1	GS/S1
Electrostatic	−46.29	−12.45	−60.73	−20.28
Pauli	116.03	38.64	129.57	51.73
Exchange	−71.91	−46.28	−60.72	−47.02
Repulsion	187.94	84.92	190.29	98.75
Polarization	−87.81	−4.52	−85.68	−11.61
Total	−18.07	21.67	−16.84	19.84

However, the most significant change is observed in the polarization energy, which increases dramatically and becomes the dominant stabilizing term. The relative increase in polarization is particularly pronounced in Ne_2, where it exceeds an order of magnitude compared to the ground-state configuration. In Kr_2, a similar trend is observed, although the relative increase is somewhat smaller due to the already significant polarization present in the ground state.

This large increase in the polarization energy in the excited state cannot be attributed solely to purely dispersion forces. Instead, it reflects the influence of the resonance character of the excited-state wavefunction, including exciton and charge-transfer resonance forms. Thus, when the wavefunction—and consequently the reference electron density defined in Eq. 2—is fully relaxed to that of the complex, the resulting polarization densities shown in Fig. 2 are obtained. For both Ne_2 and Kr_2 excimers, these densities display features that deviate from those typical of purely dispersive interactions with an accumulation and depletion of electrons in the intermolecular and intramolecular regions, respectively. This is, however, observed for the Kr_2 dimer in the GS/S1 state. Consistent with the ALMO-EDA study by Ge and Head-Gordon [11]—which identified exciton splitting as the leading contribution to the interaction energy in rare-gas excimers—our results reveal clear signatures of this effect in the polarization density. In Ne_2, the influence is particularly strong, as electron density shifts away from the bond axis into regions perpendicular to it, extending this intratomic electron redistribution into the interatomic region. This reflects the role of exciton resonance in shaping the polarization density. For Kr_2, in addition to this behavior, a faint sandwich-like redistribution appears within the interatomic region: electron density accumulates at the center while depletion occurs on either side, consistent with a small purely dispersion contribution.

These results suggest that the stabilization of rare-gas excimers is governed primarily by polarization effects arising from excitation-induced charge redistribution, rather than by dispersion alone. The strong binding is also assisted by an increase in the attractive first-order terms—namely electrostatic and intermolecular exchange interactions—which partially offset the significant rise in spin repulsion associated with the close interatomic distances in the excimers.

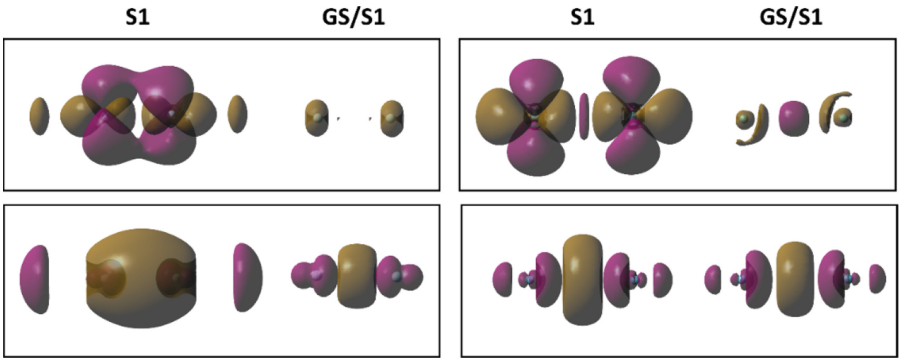


Fig. 2. Polarization (upper panels) and spin-repulsion (bottom panels) densities obtained for the lowest-lying excited state (S_1) and unrelaxed ground state (GS/S_1) complexes of Ne_2 (left panels) and Kr_2 (right panels). Brown and magenta colours represent negative and positive densities, respectively. The plots are represented using an isosurface value of $2{\cdot}10^{-3}$ au.

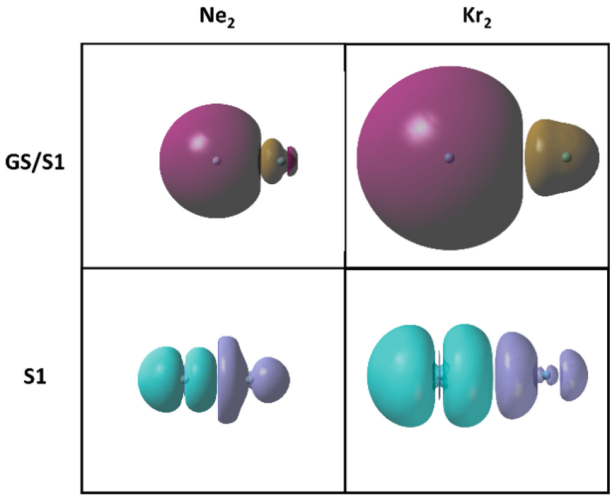


Fig. 3. DAFH density plots for the GS/S_1 states of Ne_2 and Kr_2 dimers (upper panels) and DAFH density difference upon excitation to the S_1 state (bottom panels). Magenta and violet colours represent positive density whereas brown and aquamarine blue colours represent negative density. The plots are represented using an isosurface value of $2{\cdot}10^{-3}$ au.

3.3 Intermolecular Electron Delocalization

The analysis of electron delocalization provides compelling evidence for the emergence of covalent character in rare-gas excimers strongly linked to the exciton resonance mechanism [15].

DAFH densities are depicted in Fig. 3. In order to highlight the bonding changes associated to the excitation, in Fig. 3 we have represented for the excited states the difference of the DAFH densities with respect to the unrelaxed ground states. They

show a pronounced increase in electron sharing upon excitation, with electron density associated with one atom extending significantly into the region of the other. This effect is absent in the unrelaxed ground-state configurations, where electron delocalization is negligible or even slightly antibonding.

Quantitative analysis through delocalization indices confirms this picture (see Table 2). In the excited states, DIs take large positive values, indicating strong electron sharing and a significant degree of covalent interaction. In contrast, the unrelaxed ground-state configurations exhibit small negative values, consistent with non-bonding or antibonding character.

Table 2. DIs for the Ne_2 and Kr_2 complexes in their lowest-lying excited state (S1) and in the unrelaxed ground state obtained upon vertical energy relaxation (GS/S1).

	S1		GS/S1	
	QTAIM	Mulliken	QTAIM	Mulliken
Ne-Ne	0.614	0.480	−0.013	−0.019
Kr-Kr	0.764	0.601	−0.033	−0.045

These results demonstrate that excitation induces a transition from weak, noncovalent interactions to strongly delocalized bonding characterized by significant electron sharing.

4 Conclusions

This work presents a thoroughly study of the bond strength in rare-gas excimers based on interaction energy decomposition and electron delocalization analysis. The results of the energy decomposition indicate that although vertical electronic excitation from the hypothetical geometrically unrelaxed ground-state complexes leads to a significant enhancement of all interaction energy components, polarization energy clearly emerges as the dominant stabilizing contribution. Contrary to intermolecular ground-state complexes formed by non-polar systems, the enhancement of polarization energy cannot be explained solely by an increase of dispersion forces. Conversely, the polarization density reflects the large influence of excitonic resonance on the electron charge redistribution upon complexation, in agreement with previous works where excitonic splitting was pointed out as the main stabilizing factor [11].

Both qualitative (DAFH densities) and quantitative (DIs) analyses reflect a large increase of interatomic electron delocalization upon electronic excitation, revealing a significant covalent character in the excimers. This contrasts with the weak and even antibonding character observed in the corresponding unrelaxed ground-state complexes, associated with small and negative values of DIs.

In summary, excitation-induced electron delocalization in rare-gas excimers, arising from excitonic resonance effects, produces a large enhancement of the polarization energy, largely stabilizing the complexes in the excited state. This stability is also

assisted by the increase of other attractive energy components, such as electrostatic and intermolecular exchange, which partially offset the significant enhancement of the spin repulsion energy.

The findings presented in this work provide a physically grounded picture of bonding in rare-gas excimers and contribute, in general, to our understanding of intermolecular interactions in excited state complexes.

Acknowledgments. This study was funded by Ministerio de Ciencia e Innovación (grant number PID2022-138023NB-I00) and Xunta de Galicia (grant number GRC2024/27).

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

References

1. Birks, J.B.: Excimers. Rep. Prog. Phys. **1975**(38), 903–974 (1975)
2. Birks, J.B.: Photophysics of Aromatic Molecules. Wiley, London (1970)
3. Förster, T., Kasper, K.: Ein Konzentrationsumschlag der Fluoreszenz des Pyrens. Z. Phys. Chem. **1**, 275–277 (1954)
4. Valeur, B., Berberan-Santos, M.N.: Molecular Fluorescence: Principles and Applications, 2nd edn. Wiley-VCH, Weinheim (2012)
5. Winnik, F.M.: Photophysics of preassociated pyrenes in aqueous polymer solutions and in other organized media. Chem. Rev. **93**, 587–614 (1993)
6. Turro, N.J., Ramamurthy, V., Scaiano, J.C.: Modern Molecular Photochemistry of Organic Molecules. University Science Books, Los Angeles (2010)
7. McCusker, M.: The rare gas excimers. In: Rhodes, C.K. (ed.) Excimer Lasers Topics in Applied Physics, vol. 30, pp. 47–86. Springer, Berlin, Heidelberg (1979)
8. Basting, D., Marowsky, G. (eds.): Excimer Laser Technology. Springer, Berlin, Heidelberg (2005)
9. Casals-Sainz, J.L., Jara-Cortés, J., Hernández-Trujillo, J., Guevara-Vela, J.M., Francisco, E., Martín Pendás, A.: Exotic bonding regimes uncovered in excited states. Chem. Eur. J. **25**, 12169–12179 (2019)
10. Jara-Cortés, J., Manuel Guevara-Vela, J., Martín Pendás, A., Hernández-Trujillo, J.: Chemical bonding in excited states: energy transfer and charge redistribution from a real space perspective. J. Comput. Chem. **38**, 957–970 (2017)
11. Ge, Q., Head-Gordon, M.: Energy decomposition analysis for excimers using absolutely localized molecular orbitals within time-dependent density functional theory and configuration interaction with single excitations. J. Chem. Theory Comput. **14**, 5156–5168 (2018)
12. Laschuk, E.F., Martins, M.M., Evangelisti, S.: Ab initio potentials for weakly interacting systems: homonuclear rare gas dimers. Int. J. Quantum Chem. **95**, 303–312 (2003)
13. Jones, A.C., Loeff, T.F., Heath, B.J., Goerigk, L.: Rare-gas excimers and exciplexes: surprising challenges for time-dependent density functional theory approaches. Polyhedron. **280**, 117678 (2025)
14. Morrison, A.F., You, Z.-Q., Herbert, J.M.: Ab initio implementation of the Frenkel-Davydov exciton model: a naturally parallelizable approach to computing collective excitations in crystals and aggregates. J. Chem. Theory Comput. **10**, 5366–5376 (2014)
15. Li, X., Parrish, R.M., Liu, F., Kokkila Schumacher, S.I., Martínez, T.J.: An ab initio exciton model including charge-transfer excited states. J. Chem. Theory Comput. **13**, 3493–3504 (2017)

16. Mandado, M.: Electron Density-Based Energy Decomposition Analysis: from QM to QM/MM Calculations, pp. 1–38. Springer, Berlin Heidelberg (2025)
17. Cooper, D.L., Ponec, R.: A one-electron approximation to domain-averaged Fermi hole analysis. Phys. Chem. Chem. Phys. **10**, 1319–1329 (2008)
18. Boys, S.F., Bernardi, F.: The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. Mol. Phys. **19**, 553–566 (1970)
19. Frisch, M. J., et al: Gaussian16 Revision C.01. (2016)
20. Mandado, M., Van Alsenoy, C.: EDA-NCI: Energy decomposition analysis of non-covalent interactions. https://github.com/marcos-mandado/EDA-NCI. Last Accessed 30 Mar 2025
21. Dennington, R., Keith, T.A., Millam, J.M.: GaussView, Version 6. Semichem Inc., Shawnee Mission, KS (2016)
22. Huenerbein, R., Grimme, S.: Time-dependent density functional study of excimers and exciplexes of organic molecules. Chem. Phys. **343**, 362–371 (2008)

Molecular Dynamics Simulations of Nanoemulsions: Effect of the Initial Geometry and Concentration

Jose Manuel Hermida-Ramon(✉) and Sonia Losada-Barreiro

Department of Physical Chemistry, University of Vigo, Lagoas-Marcosende s/n, 36310 Vigo, Spain
jose_hermida@uvigo.gal

Abstract. MD simulations were used to characterize the behavior of a natural antioxidant, δ-tocopherol, in a linoleic acid micelle in the presence of β-cyclodextrin as a surfactant and surrounded by water. Four different simulations were performed with different starting points and number of molecules in order to assess the importance of the initial structure and to evaluate the influence of concentration on the geometric disposition of the antioxidant with respect to the micelle. The results indicate that the surface area of the constructed oil droplet is enough to harbor the number of molecules that match the experimental molecular concentration, but that is not the case when the number increases to twice the experimental value. Also, even though the δ-tocopherol moves to the interior of the droplet across the simulations, its hydrophilic head moves out. The resulting data show the dispersive nature of the forces bonding the δ-tocopherol and β-cyclodextrin.

Keywords: Molecular Dynamics · Nanoemulsions · Micelles · Antioxidants

1 Introduction

Since the last decades of the twenty century, classical molecular dynamics (MD) has become a key methodology to address computationally the study of biomolecular systems and the characterization of the effects in surfaces. This technique provides atomic-level insights into different features of these systems, like conformational dynamics, folding pathways, and interfacial interactions. Although there is a considerable amount of studies that apply MD simulations to analyze the behavior and properties of micelles [1–6], it is scarcely used compared to the more established fields of biomolecular systems and surfaces. This disparity may be due partly to the challenges that arise when accurately modeling the dynamic amphiphilic nature of micelles. They include modelling, among others features, their size polydispersity, the shape fluctuations, and the solvent-exposed hydrophobic cores. This involves using sophisticated force fields and long simulation timescales, together with specialized protocols to achieve reproducible self-assembly and thermodynamic equilibrium. All these question have limited the use of MD simulations in micelles until recent times, when the advances in computational power and sampling techniques have boosted this use.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 291–299, 2027.
https://doi.org/10.1007/978-3-032-30539-8_22

The importance of micelles in different scientific fields is undeniable. They play a pivotal role in essential areas like drug delivery, cosmetics, food technology, environmental remediation and biochemistry. They are usually employed as nanocarriers that solubilize hydrophobic molecules in water via their hydrophobic cores and hydrophilic coronas. Antioxidants, such as polyphenols or phenolic compounds, are usually combined with micelles because they placed themselves mostly at the surfactant-water interface [7–9]. Positioned in this area they can scavenge reactive oxygen species and prevent lipid peroxidation—a process implicated in oxidative stress-related diseases or in food degradation. Although experimental techniques can been used to elucidate the general morphology of micelles and the incorporation of antioxidants, they are not able to provide dynamical information about the atomic interactions that rule the behavior of antioxidants in micelles. The outcome of this interactions comes as a result of a subtle balance of electrostatic and van der Waals forces, which is well described using MD with explicit solvent models and all-atom force fields like CHARMM [10] or AMBER [11]. Thus, for example, MD studies have shown that spontaneous micelle formation around membrane proteins or drugs occurs due to Lennard-Jones interactions and that the micelle stability can be tuned using additives or electric fields.

Despite all collected knowledge, there are still gaps about how specific antioxidants influence micelle dynamics. For example, how critical micelle concentration and aggregation numbers are modified. By performing MD simulations, structural properties, such radial distribution functions, and dynamical properties, as diffusion coefficients, can be calculated, which can shed light on the antioxidants behavior within the micelle. In this study, MD simulations are used to characterize the behavior of a natural antioxidant, δ-tocopherol, in a linoleic acid micelle in presence of β-cyclodextrin as surfactant and surrounded by water. This work is a new example of MD applied to soft colloidal systems, focusing in analyze how important is the assembly of the starting systems and the concentration of species in the final outcome of the simulations.

2 Computational Details

2.1 System Preparation

In order to represent the experimental preparation of the nanoemulsions and to evaluate the importance of the molar concentration of the involved species, two different systems and two routes to build them were adopted. Thus, the nanoemulsions were created using two different sets of molecules that differ in the their number:

1. 900 molecules of linoleic acid plus 4 molecules of δ-tocopherol and 13 molecules of β-cyclodextrin, surrounded by water molecules. This number of employed molecules is proportional to the expected number of molecules present when following the experimental molecular concentration and water/oil ratio.
2. 900 molecules of linoleic acid plus 9 molecules of δ-tocopherol plus 30 molecules of β-cyclodextrin, surrounded by water molecules. This is about twice as concentrated as the experimental ratio.

Regarding the way in which we built up the micelle that forms part of the nanoemulsion, we have adopted two strategies:

I. To use the Packmol package [12] to create a spherical oily droplet with the δ-tocopherol and the linoleic acid molecules. The δ-tocopherol molecules were placed randomly, and the linoleic acid molecules under the condition that the external surface of the oily droplet, which is in contact with water, was made by the hydrophilic group of these molecules. Then, the structure was minimized with the molecular mechanics modules of the ORCA package [13]. This package was used because of its features to perform minimizations and to include AMBER parameters. The resulting structure was introduced again in Packmol to randomly add the β-cyclodextrin molecules to the surface of the oily droplet. Finally, a new optimization of the whole system was performed.
II. The second strategy is analogous to strategy I, but the oily droplet is created only with the linoleic acid molecules. The δ-tocopherol molecules are placed on the surface of the oily droplet together with the β-cyclodextrin molecules.

Strategy I tries to simulate the conditions of an experimental procedure in which linoleic acid and δ-tocopherol are solvated together in water and the emulsion is formed, and then the β-cyclodextrin is mixed in and solvated. Strategy II represents an experimental procedure in which micelles of linoleic acid are produced first and then, δ-tocopherol and β-cyclodextrin are solvated. By combining the two sets of molecules and the two strategies to build the systems, four different models were obtained, each of them was simulated using molecular dynamics.

2.2 Molecular Dynamics Simulations

The same procedure was used to perform the simulation of all the proposed systems. Thus, the aqueous phase (solvating water) was simulated by employing the LEAP module of the AMBER suit of programs, with a TIP3P water truncated octahedron box surrounding the oily droplet so that the distance between any point of the oil droplet and any size the octahedron box was not less than 23 Å. This distance was chosen to limit the aqueous volume and get reasonable computational calculation times. A representation of the resulting model is depicted in Fig. 1. The generalized AMBER force field (gaff) parameters were used to describe the interactions between linoleic acid, δ-tocopherol and β-cyclodextrin molecules, and the TIP3P model [14] was used to describe the water molecules. These are commonly used force fields . Before starting MD simulations, the nanoemulsion geometry was minimized without restraints with the conjugate gradient method at 0 K using periodic boundary conditions in a NVT ensemble. Long-range electrostatic interactions were calculated based on the particle mesh Ewald's method with a cut-off value of 12 Å. Next, a heating procedure was performed. This procedure consist in a MD simulation that was carried out for 20 ns and where temperature was increased in a NPT ensemble from 0 K to 300 K. The time-step of this simulation was 1 fs and a Langevin thermostat collision frequency of 1.0 was used to control the temperature. The pressure was maintained at 1 atm by using the Monte-Carlo barostat with a relaxation time of 2 ps. The SHAKE algorithm [15] was applied to constrain all bond interactions involving hydrogen. The same technical specifications were used in the subsequent equilibration of the system. The simulation was performed for 80 ns in a NPT ensemble with a time-step of 2 fs. After the equilibration, a MD production simulation with no changes

in technical parameters was performed. This production run was carried out for 1 μs and the trajectory was stored every 4 ps to a total of 250000 stored snapshots. The MD simulations and the analysis of the collected snapshots were performed with the AMBER package [11]. The visualization of the structures and trajectories resulting from the MD simulations were performed with the VMD package [16].

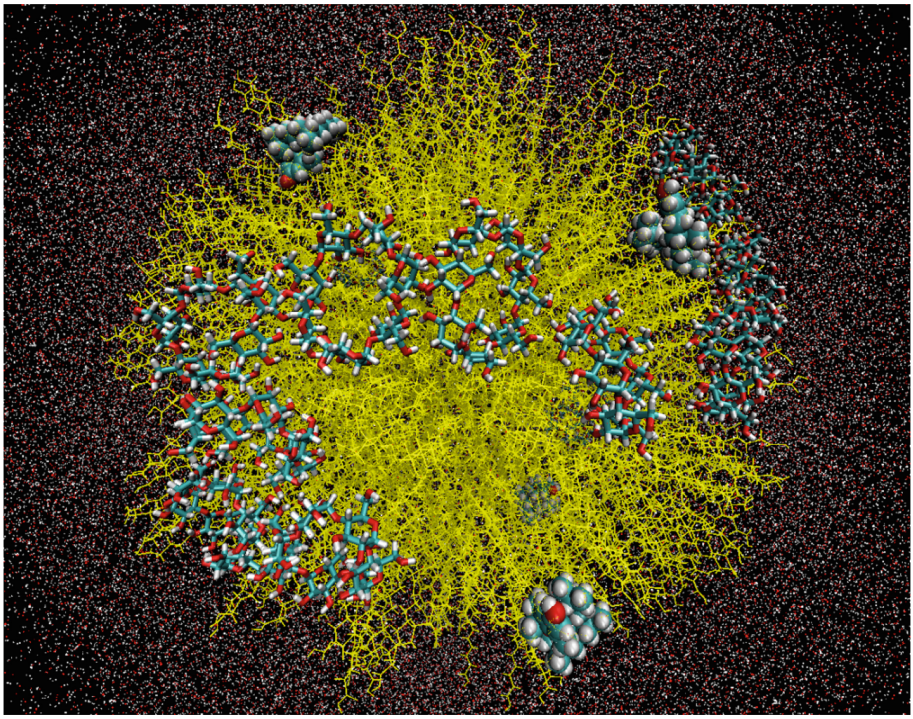


Fig. 1. Picture of one of the initial geometries employed in the molecular dynamics simulations. Water molecules are represented by points. Linoleic acid molecules are represented by lines. Van der Waals representation is used to depict δ-tocopherol molecules. β-cyclodextrin molecules are shown using a licorice representation.

3 Results and Discussion

As been already said, four simulations were performed to evaluate the influence of the starting points and the molecular concentration on the characterization of the studied system by molecular dynamics. From now on, we will label the simulation as 4Toc-I, 4Toc-II, 9Toc-I, and 9Toc-II. 4Toc corresponds to the system with four δ-tocopherol and thirteen β-cyclodextrin molecules, and 9Toc corresponds to nine δ-tocopherol and thirty β-cyclodextrin molecules. I and II correspond to strategies I and II, respectively. Thus, by comparing 4Toc-I and 9Toc-I, or 4Toc-II and 9Toc-II the effect of the molecular concentration can be analyzed. Also, if the comparison is between 4Toc-I and 4Toc-II

or between 9Toc-I and 9Toc-II, then it is the effect of the starting point of the simulation the subject that is evaluated.

Figure 2 depicts the density of probability for each of the components of the nanoemulsion. It can be seen in this figure that the thickness of the interface increases when the number of δ-tocopherol and β-cyclodextrin molecules is increased, regardless of the strategy used to build the initial structure. However, if the same number of δ-tocopherol and β-cyclodextrin molecules are considered, it can be seen that strategy II, to start the simulations with these molecules at the surface of the oil droplet, decreases the size of the interface. This may be as a consequence of δ-tocopherol and β-cyclodextrin interacting and intertwining, creating a more stable net on the surface of the droplet that produces a clearer boundary between the interface and the aqueous phase.

For this kind of nanoemulsions, it is of great interest to analyze how the δ-tocopherol is placed with respect to the surface of the droplet, since when it is localized at the interface, it can play efficiently its role as an antioxidant. Even when the starting configuration involves the antioxidant at the surface of the oil droplet, Fig. 2 shows that the δ-tocopherol moves to more profound positions across the simulation. The largest part of the tocopherol population is located within the droplet for all simulations. Yet, there are still some perceptible differences in the interface population between simulations performed with strategy I and those performed with strategy II. Although not too big, these differences may be a consequence of the intermolecular interactions between δ-tocopherol and β-cyclodextrin molecules. More precisely, due to the attractive forces between the hydrophilic head of δ-tocopherol and the hydroxyl groups of β-cyclodextrin.

When comparing the density of probability of δ-tocopherol for the 4Toc-II and 9Toc-II simulations, it can be observed that, of the two, the model with the higher molecular concentration is more populated inside the droplet. It seems to suggest that the surface area of the droplet is not enough to hold a concentration as high as that represented by model 9Toc, although from the proportions depicted in Fig. 1, it may look otherwise. This effect combines with the others already discussed, resulting in the translation of the δ-tocopherol molecules to the interior of the oil droplet.

In the previous paragraphs, the distribution of δ-tocopherol as a whole molecule has been discussed. In order to perform that, the center of mass of the molecule was used to obtain the density distribution. However, different observations can be obtained by analyzing the distribution of the hydrophilic head of δ-tocopherol instead of the whole molecule. This can be done by calculating the density of probability with respect to the oxygen atom of this hydrophilic head. Thus, when simulations with strategy I are considered, it can be seen that, even though most of the population is inside the droplet, the density presents one peak in the interface in both simulations. This may reflects the fact that, although the hydrophilic head started inside the droplet at the initial point, the simulation process moved this moiety out for some of the δ-tocopherol molecules. This behavior is opposite of what is observed for the molecule as a whole. Because the hydrophobic chain of the tocopherol gradually penetrates deeper into the micelle. Figure 2 shows that, when simulations 4Toc-II and 9Toc-II are analyzed, the largest population of the hydrophilic head is in the interface for 4Toc-II, and that 9Toc-II has a high peak at the interface, but it is counterbalanced by two peaks inside the droplet that, when added up, present a larger population than the one in the interface. This feature of

simulation 9Toc-II seems to support the observation that the surface area of the drop is not big enough for the high molecular concentration that this system represents.

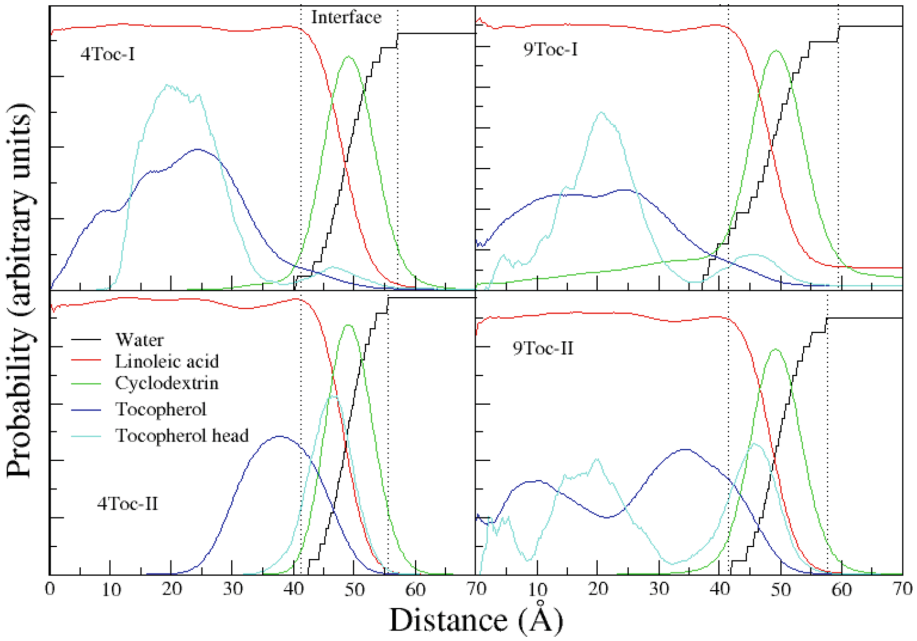


Fig. 2. Density of probability of the individual components of the nanoemulsion, plus the hydrophilic head of δ-tocopherol, with respect to the center of mass of the oil core of the nanoemulsion droplet for each of the four molecular dynamics simulations performed. The interface is marked by the dotted lines. See the text for an explanation of the labeling.

According to our results, the β-cyclodextrin molecules are placed mainly in the interface. By taking this into account, the number of hydrogen bonds that the δ-tocopherol establishes with the linoleic acid and the cyclodextrin may provide complementary information about the location of the antioxidant. Table 1 gives the percentage of hydrogen bonds formed by the hydroxyl hydrogen of one δ-tocopherol molecule with the carboxylic oxygen of the linoleic acid and any oxygen of the cyclodextrin. When considering the two simulations performed with strategy I, the results show that the one with the higher concentration has a lower percentage of hydrogen bonds with linoleic acid and a higher percentage with cyclodextrin. This observation agrees with the density curves in Fig. 2, where the peak in the interface is higher for 9Toc-I than for 4Toc-I. If simulations 4Toc-II and 9Toc-II are considered, the percentage of hydrogen bonds increases with concentration for both: linoleic acid and cyclodextrin, reflecting the shape of the curves in Fig. 2, where the tocopherol head is mainly in the interface for the 4Toc-II model but for the 9Toc-II model it has two peaks inside the droplet.

Table 1. Percentage of hydrogen bonds established by the hydrogen atom of the hydrophilic head of one δ-tocopherol molecule with the carboxyl oxygen of linoleic acid molecules and the oxygen atoms of β-cyclodextrin molecules.

	4Toc-I	9Toc-I	4Toc-II	9Toc-II
$O_{Linoleic}$	49.0%	37.4%	14.1%	23.2%
$O_{Cyclodextrin}$	0.0%	2.3%	1.9%	3.1%

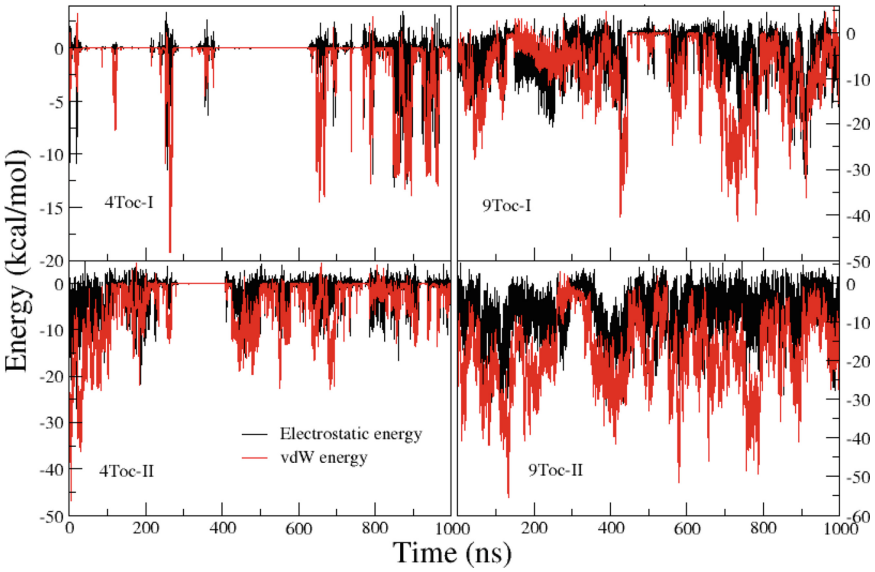


Fig. 3. Interaction energy between the δ-tocopherol and the β-cyclodextrin molecules split in the electrostatic and van der Waals terms along the simulation. See the text for an explanation of the labeling.

According to the literature [17, 18] δ-tocopherol and β-cyclodextrin can form an intermolecular complex in which the tocopherol is inserted into the cyclodextrin ring. In our simulations, we did not evaluate directly the existence of this kind of complex from a geometric point of view. Yet, we have obtained the interaction energy between tocopherol and cyclodextrin. Figure 3 presents the electrostatic and van der Waals contributions to the interaction energy between δ-tocopherol and β-cyclodextrin along the simulations. This figure shows that the higher the concentration, the stronger the interaction energy. Moreover, when using strategy II, with both groups of molecules initially on the surface, the interaction is stronger. The stronger the interaction, the more likely is that the molecules form intermolecular complexes. Although strategy II favors stronger interactions between these molecules, Fig. 3 illustrates that there are some parts of the simulation of the 4Toc-II model in which the interaction energy is zero or close to it. Suggesting that, along the simulation, the molecules move closer together, then move apart to a distance at which they cannot see each other, and then move closer again.

Finally, Fig. 3 shows that, along the simulations, the van der Waals term is usually more important to the interaction energy than the electrostatic term. This indicates that the interaction represented by these curves is more closely related to dispersive forces than to electrostatic forces.

4 Conclusions

Molecular dynamics simulations were used to characterize the behavior of a natural antioxidant, δ-tocopherol, in a linoleic acid micelle in the presence of β-cyclodextrin as a surfactant and surrounded by water. In order to check the effect of the molecular concentration and the importance of the assembly of the starting systems four different sets of simulations were performed. The strategy for building up the initial geometries was designed to represent the difference between an experimental procedure in which linoleic acid and δ-tocopherol are solvated together in water, resulting in the formation of an emulsion, and another procedure in which the emulsion is created only with linoleic acid and subsequently the δ-tocopherol is solvated.

It was found that the thickness of the interface decreases in simulations in which δ-tocopherol and β-cyclodextrin were placed on the surface of the droplet at the start of the simulation. This can be due to the two groups of molecules intertwine to create a stable network on the surface of the droplet that contributes to a more defined boundary of the interface.

Although it is not entirely clear, it seems that the behavior of the δ-tocopherol as a whole and that of its hydrophilic head are opposite to each other. Along the simulation, the δ-tocopherol appears to move toward the interior of the droplet, while the hydrophilic head moves toward the interface. This may suggest that the δ-tocopherol positions itself to interact with the cyclodextrin and the aqueous medium.

The analysis of the interaction between the cyclodextrin and the tocopherol shows that of dispersive forces predominate over electrostatic interactions. Moreover, this analysis reveals that the molecules can arrange themselves in such a way that they cannot see each other. This results in sections of the simulation in which the interaction energy is zero.

Acknowledgments. This study was funded by Ministerio de Ciencia e Innovación (grant number PID2022-136443OB-I00) and Xunta de Galicia (grant number GRC2024/27).

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

References

1. Razavi, L., Raissi, H., Farzad, F.: Validation of an MD simulation approach for electrical field responsive micelles and their application in drug delivery. Sci. Rep. **13**, 2665 (2023)
2. Bond, P.J., Cuthbertson, J.M., Deol, S.S., Sanso, M.S.P.: MD simulations of spontaneous membrane protein/detergent micelle formation. J. Am. Chem. Soc. **126**, 15948–15949 (2004)

3. Song, Y., Lee, J.H., Jung, I., Seo, B., Hwang, H.: Molecular dynamics simulations of micelle properties and behaviors of sodium lauryl ether sulfate penetrating ceramide and phospholipid bilayers. J. Phys. Chem. B. **124**, 5919–5929 (2020)
4. Cheng, X., Kim, J.-K., Kim, Y., Bowie, J.U., Im, W.: Molecular dynamics simulation strategies for protein–micelle complexes. Biochim. Biophys. Acta Biomembr. **1858**, 1566–1572 (2016)
5. Faramarzi, S. et al.: Molecular dynamics simulations as a tool for accurate determination of surfactant micelle properties. Langmuir. **33**, 9934–9943 (2017)
6. Garde, S., Yang, L., Dordick, J.S., Paulaitis, M.E.: Molecular dynamics simulation of C8E5 micelle in explicit water: structure and hydrophobic solvation thermodynamics. Mol. Phys. **100**, 2299–2306 (2002)
7. Kumar, V. et al.: Tuning cationic micelle properties with an antioxidant additive: a molecular perspective. Langmuir. **37**, 4611–4621 (2021)
8. Lü, J.-M., Lin, P.H., Yao, Q., Chen, C.: Chemical and molecular mechanisms of antioxidants: experimental approaches and model systems. J. Cell. Mol. Med. **14**, 1582–1838 (2010)
9. Chatzidaki, M.D., Papavasileiou, K.D., Papadopoulos, M.G., Xenakis, A.: Reverse micelles as antioxidant carriers: an experimental and molecular dynamics study. Langmuir. **33**, 5077–5085 (2017)
10. Vanommeslaeghe, K. et al.: CHARMM general force field (CGenFF): a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. J. Comput. Chem. **31**, 671–690 (2010)
11. Pearlman, D.A. et al.: AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Comput. Phys. Commun. **91**, 1–41 (1995)
12. Martínez, L., Andrade, R., Birgin, E.G., Martínez, J.M.: PACKMOL: a package for building initial configurations for molecular dynamics simulations. J. Comput. Chem. **30**, 2157–2164 (2009)
13. Neese, F., Wennmohs, F., Becker, U., Riplinger, C.: The ORCA quantum chemistry program package. J. Chem. Phys. **152**, 224108 (2020)
14. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L.: Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. **79**, 926–935 (1983)
15. Ryckaert, J.P., Ciccotti, G., Berendsen, H.J.C.: Numerical integration of cartesian equations of motion of a system with constraints-molecular dynamics of N-alkanes. J. Comput. Phys. **23**, 327–341 (1977)
16. Humphrey, W., Dalke, A., Schulten, K.: VMD: visual molecular dynamics. J. Mol. Graph. Model. **14**, 33–38 (1996)
17. Koontz, J.L., Marcy, J.E., O'Keefe, S.F., Duncan, S.E.: Cyclodextrin inclusion complex formation and solid-state characterization of the natural antioxidants α-tocopherol and quercetin. J. Agric. Food Chem. **57**, 1162–1171 (2009)
18. Khanittha, K., Nutho, B., Krusong, K., Poo-arporn, R.P., Rungrotmongkol, T., Hannongbua, S.: Encapsulation of α-tocopherol in large-ring cyclodextrin containing 26 α-D-glucopyranose units: a molecular dynamics study. J. Mol. Liq. **339**, 116802 (2021)

Transport Disruption Vulnerability in Europe: Who Is Worse Off?

Francesco Bruzzone[1(✉)], Hamza Begic[2], Silvio Nocera[1], and Constantinos Antoniou[2]

[1] Università Iuav di Venezia, Santa Croce 191, 30125 Venice, Italy
fbruzzone@iuav.it
[2] Technical University of Munich, Arcisstraße 21, D-80333 Munich, Germany

Abstract. Mobility inequality remains a persistent dimension of social disadvantage across Europe, yet its spatial patterns and structural determinants are insufficiently understood across scales. This study offers a dual-scale analysis of transport vulnerability, combining an EU-wide assessment with a regional focus on Italy, with particular attention to disparities between inner and non-inner areas. At the European level, vulnerability is operationalized through three key dimensions: workforce adaptability, transport system resilience, and transport affordability and choice. Using harmonized Eurostat data and the COVID-19 pandemic as a stress test, the analysis identifies a significant resilience gap across EU member states. Italy, together with eight other EU countries, falls within a car-dependent profile in which limited telework capacity and widespread car ownership mask structural fragilities, especially in peripheral areas. Zooming into this context, the Italian case reveals how car dependency translates into tangible mobility. Residents of inner areas (with limited access to essential services) exhibit significantly lower mobility dispersion, with women displaying more spatially confined movement patterns. By linking EU-level structural conditions to territorial evidence, the study contributes to a multi-scalar understanding of mobility justice and argues that effective measures, from ticketing and fares reforms to infrastructure investment, must align with local capacities to advance mobility justice, foster resilience, and ensure that no region or group remains persistently "worse off" in the face of future disruptions.

Keywords: Transport disruption · Transport vulnerability · Inner areas · Transport policy

1 Introduction

Across Europe, participation in social, economic, and civic life depends on access to safe, reliable, and affordable transport, yet this access is highly uneven across territories and social groups. Persistent socio-economic disparities, spatial inequalities, demographic change, and the increasing frequency of disruptive events have heightened concerns about how vulnerable populations cope when transport systems are disrupted [1]. In this perspective, transport disadvantage is best understood not as a simple lack of mobility supply, but as a multidimensional condition shaped by exposure to disruptions, sensitivity to their impacts, and unequal capacities to anticipate, absorb, adapt, and recover [2].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 300–310, 2027.
https://doi.org/10.1007/978-3-032-30539-8_23

Recent research highlights the close interdependence between transport vulnerability and resilience. Vulnerability captures pre-existing conditions of exposure and sensitivity, while resilience refers to the capacity of systems and users to cope with and recover from disruptive events. Groups already facing transport poverty, i.e., limited availability, accessibility, affordability, and adequacy of transport, tend to have lower preparedness and adaptability, making them less able to cope with shocks. Conversely, inclusive system design, targeted policies, and better preparedness can enhance resilience and reduce risks of long-term social exclusion [3].

Within the European Union (EU), these issues increasingly manifest along a core–periphery gradient. A substantial share of the population faces transport poverty, with significant disparities across regions, income groups, and demographic categories. Rural and peripheral areas, outermost regions, and socio-economically deprived urban neighborhoods are particularly affected, often combining limited transport provision with high exposure to climate risks and weak institutional capacity. Disruptions in such contexts can trigger cascading effects, including job loss, missed healthcare, educational disadvantages, and social isolation [4].

This paper examines how vulnerability to transport is structured across socio-spatial scales in the EU, with a particular focus on inner and peripheral areas, using Italy as a case study. Vulnerability is operationalized through three interlinked pillars: affordability (financial constraints and car dependence), adaptability (workforce capacity for remote work), and system resilience (recovery of public transport systems). Using the COVID-19 pandemic as a universal stress test, we examine how these dimensions interact across EU member states and identify where car-dependent populations in rural and inner areas are most exposed to structural weaknesses of mobility.

The study makes two main contributions. Conceptually, it advances a multi-scalar framework that links EU-level structural conditions with small-scale territorial evidence, treating transport vulnerability as a relational and context-dependent phenomenon shaped by the interaction between socio-economic characteristics and transport system configurations. The three-pillar approach, exploring adaptability, resilience, and affordability, captures labor-market, infrastructural, and financial constraints simultaneously, offering a more comprehensive perspective than single-dimension frameworks. Methodologically, the research integrates two analytical approaches that are rarely combined: (i) a macro-level statistical analysis of harmonized European datasets, enabling cross-country comparison and the conceptualization of vulnerability profiles; and (ii) a regional analysis based on mobile phone data, using an entropy index to measure the spatial dispersion of mobility with high granularity. The COVID-19 pandemic serves as a temporally consistent benchmark for assessing system responses across all EU member states. Structural vulnerabilities identified at the European scale are traced in their concrete regional manifestations, providing a replicable framework for multi-scalar analysis of transport disparity.

The remainder of the paper is structured as follows. Section 2 reviews the evolution of the concepts of transport vulnerability and resilience, with particular attention to socio-spatial inequalities. Section 3 presents the methodological framework for operationalizing vulnerability at the EU macro-level and at the regional scale, including the selection of indicators and data sources. Section 4 reports and discusses empirical results,

highlighting spatial patterns of deprivation, adaptability gaps, and resilience gaps, and identifying vulnerability hotspot profiles in inner and peripheral areas at both scales of analysis. Finally, Sect. 5 concludes by outlining policy implications for mobility justice and territorial resilience.

2 Literature Review

Research on transport vulnerability and resilience has developed along partly distinct but increasingly convergent stands. Early work in transport engineering and network science conceptualized vulnerability primarily as a system-level property, focusing on the susceptibility of networks to failures or capacity reductions and measuring impacts on performance, connectivity, and serviceability [5]. While this approach was effective in identifying critical links, bottlenecks, and recovery dynamics, it largely treated demand as homogeneous and paid limited attention to the differentiated social consequences of disruptions. In parallel, resilience was included in transport research through ecological and system-thinking perspectives. Starting with Holling's [6] work, resilience was defined as the capacity of a system to absorb disturbances, adapt to changing conditions, and maintain essential functions without shifting into an undesirable state. Subsequent distinctions between "engineering" resilience (rapid return to equilibrium) and broader socio-ecological resilience (adaptation, learning, and transformation) shifted the focus beyond technical robustness toward governance, institutional capacity, and user behavior. Within this perspective, vulnerability and resilience are increasingly understood as complementary dimensions: the former captures structural exposure and sensitivity, the latter reflects the dynamic capacity to respond and recover [2, 7].

More recent literature reframes transport vulnerability as a multidimensional and relational condition arising from the interaction between infrastructure, socio-economic characteristics, and spatial context. Disruptions do not affect all users equally; rather, their impacts are mediated by income, employment conditions, gender, age, disability, and residential location. Low-income households, for example, face affordability constraints and greater dependence on often fragile public transport systems, limiting their ability to absorb additional costs or adapt to service disruptions. In contrast, higher-income groups typically benefit from diversified mobility options and greater flexibility, including access to remote work [3]. Employment conditions and spatial context further shape vulnerability. Workers in location-dependent or low-wage occupations have limited flexibility and are more exposed to service disruptions, whereas knowledge-based workers can more easily adjust schedules or work remotely. Rural, peripheral, and low-density areas tend to have less modal diversity, fewer alternative routes, and longer distances to essential services. In such situation, even minor disruptions can produce significant accessibility losses. The notion of "forced car ownership" captures how residents may rely on private vehicles despite financial constraints due to the absence of viable alternatives, thereby increasing exposure to fuel price volatility and maintenance costs [8]. In urban settings, deprived or peripheral neighborhoods often experience lower service quality, weaker redundancy, and higher disruption risks than more central areas, reinforcing pre-existing inequalities in accessibility and safety [9].

Other socio-demographic factors reinforce these inequalities. Older adults and persons with disabilities depend on specific infrastructure and accessibility features and

may lack the capacity to reroute or switch modes during disruptions, leading to temporary or prolonged immobility. Gendered travel patterns, shaped by care responsibilities and safety concerns, make women particularly sensitive to disruptions that break trip chains or require unfamiliar routes. Migrants and ethnic minorities may face language and information barriers, especially when disruption information is not accessible or multilingual, with direct impacts on residential segregation. Across these dimensions, limited preparedness, constrained resources, and reduced access to information translate into lower adaptive capacity and resilience, even where nominal exposure to disruptions is similar [10].

Comparative research confirms that transport disruptions exacerbate pre-existing socio-spatial inequalities, with consequences for access to employment, healthcare, education, and social participation. In the EU, recent evidence highlights that transport poverty is widespread, with particularly acute problems in rural and peripheral areas, outermost regions, and socio-economically deprived urban neighborhoods. These territories often combine weak transport provision, higher exposure to climate and infrastructure risks, and limited institutional capacity. Inner and peripheral areas thus emerge as key arenas where socio-economic disadvantage intersects with fragile mobility systems.

The literature review provides three key insights for this study. First, transport vulnerability and resilience are deeply social as well as technical, and must be analyzed jointly, as interdependent dimensions linking structural conditions with dynamic responses. Second, vulnerability is inherently multi-scalar and shaped by the interaction between socio-economic characteristics and spatial context, with peripheral areas often facing compounded disadvantages in both exposure and adaptive capacity. Third, addressing transport disadvantage requires analytical frameworks and policy approaches that explicitly link equity, mobility justice, and territorial resilience. These insights motivate the analytical approach adopted in this paper, combining EU-wide assessments with regional analyses to capture how structural vulnerabilities materialize across territories.

3 Method and Data

This section translates the conceptual framework into measurable indicators at the EU and regional levels. The goal is to capture how affordability, adaptability, and system resilience combine to shape vulnerability to transport disruptions, then looking at how the analytical framework translates in people's behavior at the regional scale.

3.1 EU-Level Vulnerability Assessment

The macro-scale analysis draws on Begic et al. [11] and uses harmonized Eurostat datasets to operationalize transport vulnerability across three dimensions. Workforce adaptability is measured by the percentage of employed persons who never work from home [12]; transport system resilience by the change in public transport modal share relative to the 2019 pre-pandemic baseline [13]; and transport affordability and choice by the share of the population who cannot afford a car, complemented by the enforced lack of a car indicator [8, 14]. Coverage extends to all EU member states plus Norway and Switzerland, which are already integrated into the EU-wide statistical framework.

The COVID-19 pandemic serves as a universal stress test, providing a temporally consistent dataset that allows comparison of a pre-disruption baseline (2019), a shock phase (2020–2021), and a recovery trajectory (2022–2024). The analytical approach combines spatial mapping of socio-economic deprivation, country-by-country tracking of each vulnerability dimension, and a synthesis graph that classifies national profiles by workforce adaptability and car affordability, enabling the identification of distinct vulnerability hotspot typologies across the EU landscape, with emphasis on the core-periphery divide.

3.2 Regional-Scale Empirical Case: Friuli Venezia Giulia (FVG)

To investigate how the structural vulnerability conditions identified at EU level materialize in territorial mobility patterns, the study proposes a regional-scale analysis of FVG, in northeastern Italy, building on a recent big-data approach by Bertocchi et al. [15]. FVG is a border region with strong rural-urban and rural-touristic divide (Fig. 1).



Fig. 1. Friuli Venezia Giulia (FVG): location and inner and non-inner areas. Own elaboration on [15, 16].

The empirical base is a cell phone big data (CBD) dataset using cell-ID signal covering trips performed within FVG throughout 2023 by 86 million unique travelers across 204 analytical units (i.e., the smallest possible administrative subdivision, either a municipality or a district of a larger town). Data are disaggregated by gender and residential location, and trips are classified into five user clusters based on distance and origin-destination characteristics. These are short- and long-distance commuters (who return to the starting point daily; SC and LC), short- and long-distance travelers (who do not start and end the day at the same location, ST and LT), and Italian visitors (who enter or exit FVG from neighboring regions). To measure mobility dispersion across the region's 204 analytical units, a Shannon entropy index [17] is calculated separately by gender and traveler type (Formula 1):

$$H = -\sum_{i=1}^{n} p_i \cdot \ln(p_i) \tag{1}$$

Where:

- n is the total number of destinations;
- p_i is the proportion of trips directed to each area i relative to total trips;
- ln is the natural logarithm, used to calculate the contribution of each destination to the entropy H.

The index is then normalized enabling cross-group comparability (Formula 2):

$$\overline{H} = \frac{H}{\log(n)} \tag{2}$$

Where $log(n)$ is the maximum possible entropy if movements are evenly distributed across all 204 areas.

A key analytical distinction is drawn between inner areas (defined as municipalities with restricted access to essential services; Italian Territorial Cohesion Strategy [16]) and more connected non-inner areas, allowing the study to empirically examine how car dependency and territorial isolation compound mobility disadvantage at the sub-national scale.

4 Results

Before examining transport-specific dimensions, it is essential to establish where socio-economic vulnerability is concentrated. Analysis of material and social deprivation rates by degree of urbanization, based on [8, 18] reveals a clear and policy-relevant divide: in Eastern and Southern European countries (Romania, Bulgaria, Hungary, Greece, Slovakia, Croatia), deprivation is predominantly rural, while in many Western European countries (Belgium, Netherlands, France, Germany, Spain), it is concentrated in urban areas. This spatial pre-condition is critical because it determines how transport vulnerability manifests [19]: a deprived household in a dense Belgian city faces primarily affordability challenges within a relatively functional public transport network, whereas a deprived rural household in Romania confronts a fundamental lack of viable alternatives to the private car.

The workforce adaptability analysis confirms a deep resilience asymmetry. Countries with high adaptability (Norway, the Netherlands, and Ireland) recorded reductions of 20–35% in the share of workers who never work from home during the pandemic peak and have largely sustained these gains. In contrast, countries in Southern and Eastern Europe showed much weaker adaptation, peaking at only 7.5% for the most vulnerable group, with a notable snap-back toward pre-crisis norms by 2023. The public transport resilience dimension compounds this picture: while high-resilience countries such as Spain, Portugal, and Ireland rebounded to modal share levels more than 10% above their 2019 baseline by 2023, the low-resilience group, predominantly Eastern European and Baltic states, remained 20 to 30% below pre-pandemic ridership.

A graphical synthesis of these dimensions is shown in Fig. 2, linking car affordability and workforce adaptability based on Eurostat [12, 14]. Four national vulnerability profiles emerge. The most critical is the "vulnerable hotspot" quadrant (Bulgaria, Greece, Hungary, Latvia, Lithuania, Romania, Slovakia), where people can neither work from

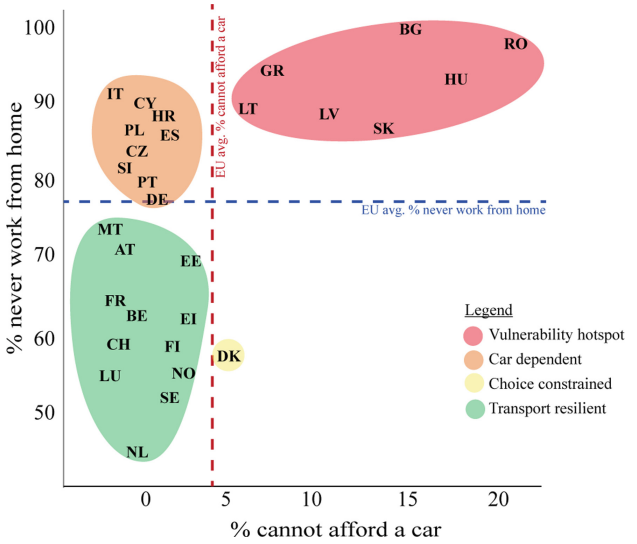


Fig. 2. Transport vulnerability, car dependency, choice constraint, and transport resilience in the EU. Own elaboration on [11].

home nor afford a car, leaving them trapped between low adaptive capacity and poorly performing public transport. Italy, together with Croatia, Germany, Poland, Portugal, and Spain, falls within the car-dependent quadrant: limited telework capacity is combined with predominantly sufficient car ownership, masking a structural fragility: households are vulnerable to fuel price shocks and any disruption that severs car-based access, precisely because no robust public transport alternative exists. This profile sets the context for the regional-scale investigation, which uses CBD to understand who moves where, and how, looking at disparities between inner and non-inner areas as well as between genders.

Across approx. 427 million trips recorded in FVG in 2023, volumes vary considerably by user type (Table 1). SC and ST generate by far the largest shares (252 and 90 million trips respectively), reflecting the dominance of routine, local mobility. LC and LT account for smaller flows, while IT exhibit, as expected, a strong seasonal pattern, peaking in August and driven primarily by tourism.

Entropy analysis reveals meaningful differences in the spatial dispersion of mobility across user groups and gender (Table 1). ST record highest entropy, reflecting a high degree of variability in destination choices. In contrast, LC show lowest entropy, indicating more spatially concentrated and predictable movements. SC score an intermediate value, with high trip volumes and limited territorial range. User type classification, then, is confirmed as a strong predictor of regional mobility behavior, showing the potential of CBD analysis in understanding mobility choices and patterns and thus effectively targeting transport policy. Gender analysis introduces a secondary but consistent modulation of this pattern: men display slightly higher entropy (i.e., spatial reach) than women across all user types, a different that becomes more pronounced in inner areas and during the winter months, suggesting that the mobility disparities in disadvantaged territories are

Table 1. Trips by user cluster and gender, and normalized entropy. Own elaboration on [15]

User cluster	Total trips	Trips by men & women (%)	Normalized Entropy Index $\overline{H}$
LT	17,701,603	59–41	0.0226
ST	90,121,171	56–44	0.0234
IT	19,284,705	60–40	0.0226
SC	252,323,371	55–45	0.0225
LC	47,406,740	56–44	0.0201
Total	426,837,590	56–44	0.0222

even more pronounced for women, worsening gendered constraints on mobility range and opportunity access, in line with findings of recent literature [20, 21].

The contrast between inner and non-inner areas provides other significant findings. In addition to exacerbating the gender gap, inner areas (82 small municipalities in FVG) exhibit substantially and statistically significant lower entropy for all users, throughout the year. This gap is most pronounced in the winter months, when the absence of leisure-driven travel removes the partial compensation that seasonal tourism provides during summer. Read in conjunction with the EU-level finding that Italy falls within a car-dependent vulnerability profile, these results indicate that territorial isolation in inner areas translates into a concrete and measurable reduction in mobility range, even more so for women and for residents of inner areas. This condition, reflecting the absence of viable public transport alternatives, carries direct implications for access to employment, healthcare, and essential services.

5 Conclusions

This study has examined how vulnerability to transport disruptions and mobility disparity are socially and spatially structured across spatial scales, from the EU-wide distribution of socio-economic exposure and adaptive capacity to the concrete patterns of territorial mobility in a car-dependent regional context. By critically elaborating on a macro-level statistical framework [11] with a regional big data analysis [15], it has shown that vulnerability is neither random nor uniform, but follows identifiable structural patterns shaped by the interaction of transport system profiles, labor market characteristics, and territorial conditions.

At the European scale, the findings confirm a persistent and geographically patterned resilience gap. The "non-teleworkable, car-constrained" hotspot, concentrated in Southeastern Europe and the Baltic states, represents a dual bind in which individuals lack both the crucial option of remote work and access to a recovering public transport system. Italy and comparable car-dependent countries occupy a structurally distinct but equally fragile position: widespread car ownership masks an underlying vulnerability to any disruption that raises the cost or feasibility of private car use, without a resilient public transport network available as a substitute. The same disruptive event therefore

produces fundamentally different consequences depending on whether populations have modal redundancy, digital work flexibility, and system capacity to absorb shocks.

At the regional scale, the FVG analysis gives empirical substance to these structural observations. Traveler type emerges as the primary determinant of mobility dispersion, with short-distance commuters (SC) and intra-province travelers (ST) generating the largest volumes of routine, local trips, while long-distance travelers (LT) exhibit more concentrated spatial patterns. The most consistent and statistically significant divide, however, runs between inner and non-inner areas: residents and travelers originating in inner territories display systematically more confined mobility throughout the year, a gap that widens for women and in winter.

Together, these findings demonstrate that transport disadvantage cannot be treated as a universal condition tractable with uniform solutions. Indeed, it requires differentiated, context-calibrated policy responses that align financial instruments, labor market flexibility, digital inclusion, and transport investment with the specific structural profile of each territory. Strengthening transport resilience should be framed as a public health, social inclusion, and socio-economic stability strategy, capable of reducing cascading risks such as income loss, missed healthcare, and long-term exclusion from civic life.

Despite allowing for interesting and high-value policy insights, this research has some limitations. At the EU scale, the analysis relies on macro-level national indicators that, while enabling consistent cross-country comparison, necessarily mask within-country heterogeneity and sub-national variation in vulnerability conditions. At the regional scale, the cell phone big data (CBD) approach captures spatial mobility patterns with high granularity but does not provide information on trip purpose or transport mode, limiting the extent to which observed differences in mobility dispersion can be directly attributed to specific structural constraints such as car dependency or public transport absence. Across both scales, the descriptive and typological nature of the analysis means that causal mechanisms behind observed patterns cannot be fully established. Future research could bridge these levels more explicitly by integrating micro-data, applying multilevel modelling or causal inference methods to test which socio-economic characteristics and system features jointly produce or sustain transport vulnerability. By identifying where vulnerability hotspots are located and how they manifest across scales, from structural resilience gaps at EU level to mobility disparities in inner areas, this paper provided a system-level basis for designing integrated policies that support territorial resilience and ensure that no region or group remains persistently worse off in the face of future disruptions or lasting disparities.

Acknowledgments. This research was supported by Horizon Europe research and innovation programme under the project SCUDO (*Safe and Comprehensive Urban Development and Transport Management for Disruption-Resilient Operations).*

Disclosure of Interests. The authors have no competing interests to declare.

References

1. Pregnolato, M., Dawson, D.: Adaptation investments for transport resilience: trends and recommendations. Int. J. Saf. Secur. Eng. **8**, 515–527 (2018). https://doi.org/10.2495/SAFE-V8-N4-515-527

2. Reggiani, A., Nijkamp, P., Lanzi, D.: Transport resilience and vulnerability: the role of connectivity. Transp. Res. A Policy Pract. **81**, 4–15 (2015). https://doi.org/10.1016/j.tra.2014.12.012
3. Martens, K.: Transport Justice: Designing Fair Transportation Systems. Routledge/Taylor & Francis Group, New York (2017). https://doi.org/10.4324/9781315746852
4. Carvalho, T., El-Geneidy, A.: Measuring the impacts of a major metro disruption in Montréal, Canada, on riders' satisfaction and willingness to recommend the service to others. Case Stud. Transp. Policy. **20**, 101480 (2025). https://doi.org/10.1016/j.cstp.2025.101480
5. Sun, D.J., Zhao, Y., Lu, Q.: Vulnerability analysis of urban rail transit networks: a case study of Shanghai, China. Sustainability. **7**, 6919–6936 (2015). https://doi.org/10.3390/su7066919
6. Holling, C.S.: Resilience and stability of ecological systems. Annu. Rev. Ecol. Syst. **4**, 1–23. http://www.jstor.org/stable/2096802 (1973)
7. Seeliger, L., Turok, I.: Towards sustainable cities: extending resilience with insights from vulnerability and transition theory. Sustainability. **5**, 2108–2128 (2013). https://doi.org/10.3390/su5052108
8. European Commission: Transport poverty: Definitions, indicators, determinants, and mitigation strategies – Final Report. Report. Publications Office of the European Union (2024). https://doi.org/10.2767/0662480
9. Lucas, K., Mattioli, G., Verlinghieri, E., Guzman, A.: Transport poverty and its adverse social consequences. Proc. Inst. Civ. Eng. Transp. **169**, 353–365 (2016). https://doi.org/10.1680/jtran.15.00073
10. Papangelis, K., Velaga, N.R., Ashmore, F., Sripada, S., Nelson, J.D., Beecroft, M.: Exploring the rural passenger experience, information needs and decision making during public transport disruption. Res. Transp. Bus. Manag. **18**, 57–69 (2016). https://doi.org/10.1016/j.rtbm.2016.01.002
11. Begic, H., Bruzzone, F., Nocera, S., Antoniou, C.: Analysis of socio-economic vulnerability to transport disruptions in the European Union. J. Urban Mobil. **9**, 100235. Manuscript under review (2026)
12. Eurostat: Employed persons working from home by professional status- % of total employment. (2024a). URL: https://ec.europa.eu/eurostat/databrowser/view/lfsa_ehomp/default. Last updated: September 2025. Database. Accessed 12 Nov 2025
13. Eurostat: Modal split of inland passenger transport. (2024b) URL: https://ec.europa.eu/eurostat/databrowser/view/tran_hv_psmod/default. Last updated: June 2025. Database. Accessed 12 Nov 2025
14. Eurostat: People who cannot afford a car, categorized by gender, age, and degree of urbanization. (2024c). URL: https://ec.europa.eu/eurostat/databrowser/product/view/ilc_mddu05a/default. Last updated: October 2025. Database. Accessed 12 Nov 2025
15. Bertocchi, D., Bruzzone, F., Nocera, S.: A big data analysis of the gender gap in mobility at the regional scale: insights from northern Italy. J. Transp. Geogr. **128**, 104390 (2025). https://doi.org/10.1016/j.jtrangeo.2025.104390(2025)
16. Presidency of the Council of Ministries of Italy: A Strategy for Inner Areas in Italy: Definition, Objectives, Tools and Governance. Materiali Uval Series. URL: https://politichecoesione.governo.it/it/politica-di-coesione/strategie-tematiche-e-territoriali/strategie-territoriali/strategia-nazionale-aree-interne-snai/. Accessed 12 May 2026
17. Poliziani, C., Rupi, F., Schweizer, J., Postorino, M.N., Nocera, S.: Modeling cyclist behavior using entropy and GPS data. Int. J. Sustain. Transp. **17**(6), 639–648 (2023). https://doi.org/10.1080/15568318.2022.2079446
18. Eurostat: Rate of material and social deprivation by degree of urbanization. (2024d). URL: https://ec.europa.eu/eurostat/databrowser/product/view/ilc_mdsd09. Last updated: October 2025. Database. Accessed 12 Nov 2025

19. Eurofound: Access to essential services for people on low incomes: energy, public transport and digital communications. Report. Publications Office of the European Union (2022). https://doi.org/10.2806/87418
20. Gauvin, L. et al.: Gender gaps in urban mobility. Human. Soc. Sci. Commun. **7**(1), 1–13 (2020). https://doi.org/10.1057/s41599-020-0500-x
21. Liu, Z., Zhao, P., Liu, Q., He, Z., Kang, T.: Uncovering spatial and social gaps in rural mobility via mobile phone big data. Sci. Rep. **13**(1), 6469 (2023). https://doi.org/10.1038/s41598-023-33123-0

Backlog Traps, Congestion and Sustainable Transition in Port Systems: A System Dynamics Approach

Ani Simonyan(✉) and Giuseppe Borruso

Department of Economics, Business, Mathematics and Statistics "Bruno de Finetti", University of Trieste, Trieste, Italy
ANI.SIMONYAN@phd.units.it, giuseppe.borruso@deams.units.it

Abstract. Ports are simultaneously experiencing rising cargo flows, system congestion, and more stringent environmental standards. This article proposes examining the transition to sustainable ports from a system dynamics perspective, using the Backlog Traps model to demonstrate why operators' and ports' frequency hinder the circular economy and investments in decarbonization initiatives, despite their long-term benefits. The task backlog is the outcome of compounded operational and environmental demands, such as delays, Carbon emissions, congestion, and waste, that diminish the resources at hand for long-term improvements. In transport hubs such as seaports and dry ports, backlog traps degrade transport efficiency and generate congestion, which, in turn, has important side effects, including longer transit times and increased CO2 emissions, thereby indirectly worsening environmental quality. In the context of this framework, the aim is to improve efficiency optimization as a key mechanism for reducing emissions and environmental impacts in congested contexts. Circular economy solutions are considered to the extent that they contribute to reducing the carbon footprint of port and hinterland operations.

Keywords: Backlog Traps · Congestion and the Sustainable transition · A system dynamics approach

1 Introduction

In recent decades, Port Systems (PS) have evolved into complex socio-technical systems integral to global logistics and supply chain networks. This transformation is primarily attributed to rising cargo volumes, expanded hinterland connectivity, and the increasing integration of maritime transport within global production systems. Concurrently, ports face mounting pressure to comply with stringent environmental regulations and support broader sustainability and circular economy goals [1].

Although the necessity for sustainable port systems is widely acknowledged, implementation remains inconsistent and frequently delayed. Conventional explanations focus on financial, technological, and regulatory challenges. However, these factors do not

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 311–326, 2027.
https://doi.org/10.1007/978-3-032-30539-8_24

fully account for the internal dynamics influencing decision-making processes within port systems.

During their day-to-day operations port authorities and port operators often face pressures due to congestion, delays and capacity constraints. Consequently, short-term solutions relegating long-term strategic investments and resources are derived from structural improvements, hindering sustainable transformation.

This paper suggests that such dynamics can be better understood within the framework of the backlog traps model, grounded in System Dynamics theory [2, 3], as it enables the analysis of feedback loops, time delays and non-linear interactions present in complex systems. It also offers a framework for understanding how operational pressures, investment behavior and system performance evolve together over time.

The Theoretical Framework introduces System Dynamics and clarifies why feedback structures and temporal delays are fundamental for interpreting port congestion. The Methodology formalizes the Backlog Traps model and incorporates multi-actor port behavior into the backlog capacity interaction. The Causal Loop Analysis and Case Study examine the reinforcing and balancing mechanisms at play, with the Trieste case demonstrating how operational and governance constraints perpetuate congestion. The Discussion unpacks these dynamics and draws attention the systemic conditions that allow backlog traps to persist. The Conclusion then reflects on how more coordinated governance, timely investment, and efficiency-oriented interventions can help steer port systems toward more sustainable trajectory.

2 Theoretical Framework

System Dynamics methodology developed in the mid-1950s at the Massachusetts Institute of Technology (MIT) by Jay W. Forrester [4], is the methodology that enables understanding and modelling the behavior of complex systems, such as ports, logistic networks and environmental transition. It offers an approach to describe the structural features of feedback loops, delays, congestion and nonlinear relationship that are typical of real-world systems. Forester's essential works Industrial Dynamics [4], Urban Dynamics [5] and World Dynamics [6], revealed that complex systems operate paradoxically and that intervention in them without considering the structures of feedback often produces unforeseen outcomes. The methodology found international application in the publication of the book "The Limits to Growth" (Club of Rome) by co-authors Donella Meadows, Dennis Meadows, Jørgen Randers and William Behrens [7], using the World3 model. The study showed how population, industrial production, resource use and pollution interact with each other through reinforcing and balancing feedback loops. Later works Limits to Growth: The 30-Year Update [8] and Randers' 2052: A Global Forecast for the Next Forty Years [9] reinforced the role of SD in long term sustainability analysis as a tool of studying environmental boundaries, resource dynamics and global transitions.

Later, John Sterman, director of the MIT System Dynamics Group, extended the methodology to the fields of organizational learning, climate policy and operational management, his textbook Business Dynamics [10] is now the global standard for system dynamics modelling. Donella Meadows developed system thinking as practical thinking

and Peter Senge introduced system dynamic into organizational strategy through his book The Fifth Discipline. Barry Richmond made dynamic modelling accessible to a non-technical audience by developing the STELLA/I Think software environment. Russell Ackoff, though working more broadly in systems science, contributed essential insights on holistic problem solving and the management of "messes" complex, interconnected challenges that cannot be decomposed into isolated parts.

2.1 System Dynamics in Transport, Logistics and Sustainability

SD has also been widely applied in transport, logistics, port operations and environmental sustainability. For the formation of congestion, infrastructural bottlenecks, investment delays, for the analysis of traffic formation, infrastructural bottlenecks, investment delays, modal shifts, environmental impact phenomena, which are characteristic of the problems of the backlog traps and sustainability challenges faced by ports and inland terminals.

H. David Ford, Khalid Saeed, and Yaman Barlas were among the first to apply SD to infrastructure investment and congestion dynamics. Ford's work on project management and capacity expansion [11] highlighted how delays in investment decisions can create reinforcing loops of congestion. Saeed's research on policy resistance [12] demonstrated that well-intentioned interventions often fail when underlying feedback structures are ignored. Barlas contributed methodological advances in model validation and sensitivity analysis [13], providing tools to rigorously test dynamic hypotheses an essential step when modelling port capacity and backlog accumulation.

Navid Ghaffarzadegan and Hazhir Rahmandad expanded SD into the study of capacity constraints, resource allocation, and delay driven instabilities. Their work on workforce bottlenecks, service delays and dynamic resource misallocation offers direct parallels to port operations, where limited birth capacity, crane availability, and hinterland connection can trigger cascading congestion [14, 15].

Lars-Göran Mattsson and Jonas Eliasson applied Sd principles to transportation system, focusing on model shifts, congestion feedback and long-term infrastructure planning. Their work demonstrated that transport networks behave as tightly coupled systems where congestion in one node propagates through the entire network an insight critical for multimodal port hubs [16].

Ruth Madlener and Reinhard Madlener contributed significantly to SD applications in energy transitions and environmental policy. Their models of emission dynamics, technology adaption and policy feedback provide a conceptual foundation analyzing environmental footprint of port operation [17, 18].

Giuliano, O'Kelly and Rodrigue although not always explicitly using SD produced influential research on port logistics, hinterland capacity and terminal capacity that align closely with SD concepts. Their findings support the SD prospective that ports must be understood as nodes embedded in broader logistical and spatial systems with feedback driven behavior [19–21].

Bichou and Gray advanced systems thinking in port performance, resilience and supply chain integration. Their work shows that port performance emerges from the interaction of multiple subsystems terminal operations, hinterland logistics, regulatory frameworks and market dynamics [22, 23].

The research of Acciaro, Gritsenko and Parola dedicated to the study of sustainability, environmental management and strategic transition in ports, emphasizes that sustainability in ports is determined by delayed investments, through regulatory responses, technological expansion cycles and behavioral responses [24–26]. These characteristics make SD particularly suitable for the structural peculiarities of ports, where such dynamics are ubiquitous:

- Backlog traps and reinforcing congestion loops,
- Delays in capacity expansion and investment responsiveness,
- Nonlinear relationships between throughput, emissions and operational efficiency,
- Policy feedback shaping long-term sustainability trajectories,
- Interdependencies between port operations, hinterland logistics and environmental outcomes.

For these reasons, SD provides a robust methodological foundation for analyzing backlog traps and designing policies that support a sustainable transition in port systems. Nonetheless, an element to be considered is the multi-layered, multi-scalar nature of ports. On one side in fact ports operate as interfaces among long-range networks, therefore as nodes in a regional or global set of connections; on the other side, ports host a set of complex activities within different and specialized terminals.

The nature of ports makes them interface between sea and land, therefore implying a set of interactions on the seaside with the maritime routes, and with the land connections. Very often they are inserted or close to urban areas, igniting process of interaction and possibly friction with the urban fabric. Port operations and activities can be therefore related to sea-side operations and external operations in land connections. Different agents are therefore involved in such a framework. This multi-institutional environment includes public port authorities, who act as cluster managers regulating infrastructure, and private terminal operators focused on maximizing operational efficiency. Effectiveness is further shaped by the heterogeneous expectations of users, including shipping lines, cargo owners, and 3PL/4PL providers. Moreover, the essential role of dock labor and the integration of inland transport providers create a performance continuum spanning maritime, terminal, and hinterland domains [27, 28].

More specifically, delving into the in-port operations, different sets of agents and activities can be highlighted: the agents and subjects directly and specifically involved in terminal activities can be categorized into management and ownership entities, the operational workforce, logistics intermediaries, and technical service providers. These actors operate within the maritime-land interface, facilitating the transfer of cargo and passengers.

Management and Ownership Entities include Port Authorities, as the primary owners and operators, most of them organized under the landlord model, acting as “cluster managers”, regulating infrastructure, managing maritime access (like dredging and navigation aids), and overseeing the relationships with terminal customers.

Terminal Operators (Port Holdings): These are often private entities that lease terminal facilities through concession agreements. They are responsible for day-to-day operations, investing in superstructure (such as cranes and yard equipment), and managing the specific cargo type the terminal is designed for (e.g., containers, liquid bulk, or dry bulk).

The Operational Workforce (Dock Labor). These are dockworkers - a fundamental production factor for terminal performance including specialized Operatives, as crane operators (or winchmen), straddle carrier/forklift drivers, and signalmen (hatch tenders) who supervise cargo movement, Lashers and Tallymen: securing and releasing cargo on ships, while tallymen (checkers) keep records of the weight and quantity of goods and check for damage. Are included supervisory Staff, foremen and supervisors managing specific groups of workers, typically assigned to one ship's hatch or shoreside crane. Technical Support: Mechanics and electricians responsible for maintaining terminal equipment. Administrative and White-Collar Staff are involved in management, information technology (IT), marketing, and legal matters.

Terminal Users and Decision-Makers include Shipping Lines (Carriers), as the primary buyers of terminal services, providing the vessels that require berthing and cargo handling. Some lines are vertically integrated, operating their own terminals (e.g., MSc). Part of this set are Cargo Owners (Shippers): Including manufacturers, retailers, and exporters/importers who determine the flow of goods. Their satisfaction is a key measure of terminal effectiveness. These are integrated by 3PL (Third Party) and 4PL (Fourth Party) Providers acting as supply chain integrators, often managing the outsourcing of transportation, warehousing, and distribution on behalf of shippers.

Technical and Intermediary Support Agents involve Nautical Service Providers, as pilots and tug operators who ensure the safe navigation of ships through access channels and into berthing basins; Intermediary Agents, as Freight forwarders and ship agents handle the booking, consolidation (like Less than Container Load cargo), and tracking of shipments; Regulatory Agents: Customs officials conduct inspections and process documentation, directly impacting "mean-time customs clearance" and terminal fluidity.

Inland Integration Agents are the fundamental agents allowing the port-region/global land connectivity, by means of Trucking Companies and Rail Operators: These subjects are vital for the hinterland operations of the terminal, providing the land-based connection to markets. Efficient interaction between terminal gate systems and these operators is critical to avoiding bottlenecks.

Research on port economics points out the ongoing tension between operational efficiency, expanding capacity and long-term strategic growth. Challenges in planning and forecasting also affect the efficiency with which port systems develop [29]. The Backlog Traps concept shows that when operational and environmental pressures build up without being addressed, it becomes harder for a system to invest in long-term improvements.

3 Methodology

The backlog traps model introduced here offers a dynamic representation of how operational pressures, investment behavior and environmental constraints interact over time to influence port performance. At the core of the methodology is a System Dynamics model structured around interacting stocks, flows and feedback loops. Two state variables backlog and capacity capture the evolution of port performance. Backlog represents the accumulation of operational and environmental pressures within the system, while capacity reflects the system's ability to process these pressures. Their interaction is formalized through a set of differential equations.

Backlog B (t) evolves according to the difference between operational pressure and system capacity:

$$\frac{dB(t)}{dt} = P(t) - C(t)$$

Where P (t) denotes demand driven operational pressure and C (t) represents the system's effective capacity. When capacity exceeds pressure, backlog decline, when operational pressure exceeds capacity, backlog grows. This formulation captures the fundamental dynamic underlying congestion formation.

Capacity C (t) evolves as a function of investment and congestion effects:

$$\frac{dC(t)}{dt} = \alpha I(t) - \beta B(t)$$

Here, α indicates the effectiveness of investment in increasing capacity, whereas β captures the negative impact of congestion on system performance. This equation embodies a duality in the evolution of capacity: investment expands capacity, while congestion erodes it.

Investment behavior is modelled as a decreasing function of backlog:

$$I(t) = I_{\max} - \gamma B(t)$$

This behavioral rule reflects decisionmaker's tendency to prioritize immediate operational needs during in heigh - pressure situations. As backlog of deferred tasks grows, attention and resources shift toward short-term firefighting, reducing the capacity to invest in long-term improvements. This mechanism is central to the formation of backlog traps.

Capacity refers not only to physical infrastructure but also to effective capacity. The latter is formed through coordination, planning and digitalization and enables capacity to be increased without the need for a proportional physical expansion and is particularly important for ports with spatial or regulatory constraints [1].

The model also incorporates the multi-actor nature of port systems. Ports operate as socio-technical environments where performance emerges from the interaction of diverse agents. Port authorities steer long term development by shaping strategic planning, infrastructure expansion and regulatory oversight, thereby influencing the direction and the timing of major investments decisions. Terminal operators oversee day-to-day cargo handling, equipment deployment and workforce scheduling, directly affecting operational efficiency and congestion levels. Hinterland operators particularly rail carriers in the case of Trieste determine the speed and reliability of inland flows, with rail bottlenecks often becoming critical constraints. Shipping lines influence peak demand through vessel scheduling and service frequency, while customs and regulatory agencies affect clearance times and administrative delays. Logistics providers and freight forewords further shape demand patterns and routing decisions. These mechanisms jointly shape how backlog accumulates, how capacity is utilized and how responsive the system is to new investments, making a dynamic modelling approach particularly suitable for capturing the complexity of port systems.

To contextualize the model, it is essential to outline the operational structure of a standard container terminal. Terminal operations unfold through a sequence of independent processes, beginning with vessel arrival and berthing, followed by quay crane operations, yard storage and handling, and finally gate or rail dispatch. Each stage introduces potential bottlenecks at the berth, within yard operations, in rail dispatching, and across administrative procedures. These constraints interact along the entire operational chain, and their system-wide implications are formalized in Section 4 through the reinforcing congestion loop.

Beyond rail-related constraints, hinterland congestion is also shaped by truck arrival patterns and parking availability. Evidence from agent-based modelling shows that inadequate truck parking facilities and uncoordinated gate arrivals can amplify queuing, increase dwell times, and propagate congestion across the terminal–hinterland interface. Such findings, including those demonstrated in the study "Reducing congestion in port hinterlands: an agent-based modelling case study of truck parking facility impacts," highlight how small-scale behavioral frictions can escalate into system-wide backlog accumulation. Integrating these dynamics into the Backlog Traps framework underscores the importance of accounting for both modal diversity and behavioral variation when analyzing port performance.

Taken together, the methodology offers a dynamic and behaviorally informed lens of examining how operational pressures, investment decisions and environmental constraints interact to shape system outcomes. By capturing these interdependences, the Backlog Traps model provides a novel perspective for understanding the persistent challenges of congestion and delayed sustainability transition in port systems.

4 Causal Loop Analysis

This section develops the causal logic underlying backlog formation and congestion dynamics in port systems. Building on system dynamics theory, the analysis focuses on how reinforcing and balancing mechanisms interact to shape operational performance. Two causal loop diagrams (CLDs) are presented to illustrate the general backlog trap and its specific manifestation in the Port of Trieste.

Figure 1 illustrates the basic feedback structure governing congestion dynamics in port systems. The diagram illustrates one reinforcing loop (R1) and two balancing loops (B1 and B2) which together shape the dynamic evolution of backlog and capacity over time. The reinforcing loop R1 captures the self-sustaining congestion mechanism: as backlogs increases, congestion intencifies, reducing transport efficiency and slowing processing rates, which in turn further amplifies backlog growth. This dynamic reflects the endogenous nature of congestion formation, in which operational delays and resource constraints reinforce one other through a positive feedback loop.

The balancing loop B1 depicts the investment-driven response. Growing demand puts pressure on capacity expansion, encouraging investments that increase physical or infrastructural capacity. Increased capacity eases congestion and gradually reduces backlog, creating negative feedback that stabilizes system performance. However, when investment decisions are delayed or insufficient, the balancing effect weakens, and the reinforcing loop to dominates.

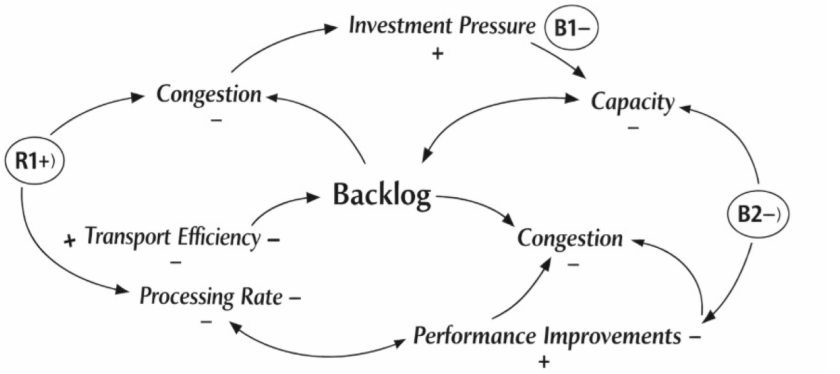


Fig. 1. General Backlog Trap: Reinforcing and Balancing Feedback Loops

The balancing loop B2 represents performance-oriented responses. Operational improvements such as process optimization, digitalization, and coordination raise effective capacity without requiring proportional physical expansion. Enhanced efficiency counteracts backlog growth, forming a second stabilizing mechanism [30, 31].

Together, the three loops describe the dynamic interaction between short-term operational pressures and long-term adaptive responses. The backlog trap arises when the reinforcing loop (R1) operates faster than the balancing loops (B1 and B2), resulting in a persistent congestion and delayed sustainability transition. This structure underscores the importance of timely investments and continuous performance improvements as complementary strategies for achieving resilient and sustainable port operations.

4.1 Trieste Specific Backlog Trap

Figure 2 adapts the general backlog trap to the specific operational and governance context of the Port of Trieste. Trieste's dependence on rail makes it sensitive to disruptions, a dynamic detailed further in Sect. 5 [32].This operational chain reinforces backlog through the same mechanism as R1 in Fig. 1 The same balancing mechanisms (B1/B2) apply, though their effectiveness is constrained by governance complexity.

Figure 2 presents the backlog trap as it manifests in the Port of Trieste, emphasizing the port's structural dependence on rail-based hinterland connectivity and the governance complexities that shape its operational performance.

The diagram extends the general backlog trap by specifying the causal mechanisms through which rail disruptions propagate across terminal operations, generating a self-reinforcing congestion spiral. The reinforcing loop R1 captures this dynamic: rail disruptions increase yard occupancy, which reduces quay crane productivity and prolongs vessel turnaround times. Longer turnaround times contribute to backlog accumulation, which further intensifies congestion and exacerbates rail disruptions. This loop reflects the tightly coupled nature of Trieste's intermodal chain, where disturbances in rail evacuation rapidly translate into system-wide congestion.

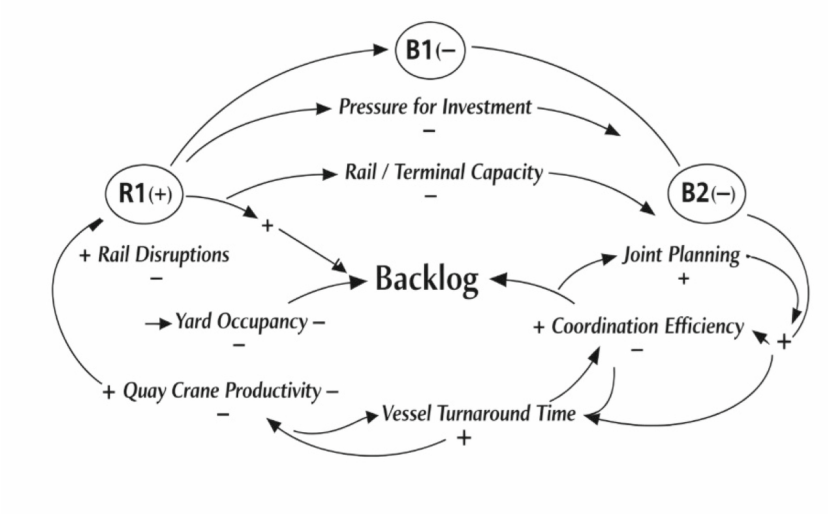


Fig. 2. Trieste-Specific Backlog Trap: Reinforcing and Balancing Feedback Loops

Two balancing mechanisms counteract this reinforcing dynamic. The balancing loop B1 represents investment-driven responses aimed at expanding rail and terminal capacity. As backlog increases, pressure to expend infrastructural improvements intensifies, promoting measures such as additional rail sidings, upgraded intermodal yards, or modernized terminal equipment. Enhancing capacity alleviates yard congestion and mitigates bottlenecks, thereby contributing to reduction of backlog. However, in Trieste, implementation of such investmments is frequently slowed by multilevel governance structures, regulatory requirements, and coordination challenges across national and transitional rail systems. These delays undermine the stabilizing effect of B1, allowing the reinforcing link to dominate during peak pressure periods.

The balancing loop B2 captures coordination-based responses. Growing backlog stimulates joint planning efforts among port authorities, terminal operators, rail carriers, and customs agencies. Improved coordination enhances scheduling efficiency, synchronizes train routes, and reduces administrative delays, thereby increasing operational capacity and curbing backlog growth. This loop highlights the importance of organizational and procedural improvements, which can be deployed more rapidly than physical investments.

Together, the three loops illustrate how Trieste's congestion dynamics emerge from the interaction of operational dependencies, infrastructural constraints, and governance structures. Under rail disruption conditions, the dominance of the reinforcing loop R1 explains the port's vulnerability to systemic congestion, while the delayed or partial activation of balancing loops B1 and B2 explains the persistence of backlog traps. This structure highlights the need for integrated investment strategies and strengthened multi-actor coordination to enhance the port's resilience and support sustainable transition.

Figures 1 and 2 establish the causal logic of backlog formation and its specific manifestation in Trieste. However, these diagrams focus primarily on feedback mechanisms.

To fully understand how these loops emerge from the underlying structure of port operations and governance, Sect. 5 introduces Fig. 3, an integrated model that synthesizes operational processes, actor roles, and feedback dynamics into a unified framework.

5 Case Study: Port of Trieste

The Port of Trieste provides an ideal context for applying the causal loop framework developed in Sect. 4. As a rail-oriented gateway on the TEN-T Baltic-Adriatic and Mediterranean Corridors, Trieste's performance is shaped by the interaction of operational constraints, governance structures, and systemic feedback mechanisms. The port of Trieste is the Northern port in the Adriatic Sea; it is the first Italian port for tonnage handled and is located at the border with Slovenia. It is characterized by its role as an international and free port, thanks to a privileged regime inherited from the past belonging to the Austro-Hungarian Empire and supported by an important rail network that allow a high percentage of transit commercial traffic to be travel by train towards inland European destinations. The port is also physically interconnected to the city of Trieste, as well as characterized by geographical constraints as the need to overcome the Karst Plateau to be connected to the national and international rail and road network. To represent these interconnections, Fig. 3 presents an integrated model that combines the operational chain, stakeholder management, and feedback loops, which together determine port efficiency.

Figure 3 presents an extended causal structure that captures the interaction between operational constraints and institutional dynamics in shaping port efficiency. The diagram depicts the systemic feedback underlying congestion in the Port of Trieste, emphasizing how physical bottlenecks and governance delays interact to create self-perpetuating inefficiencies.

At the core of the model lies the reinforcing loop R1, which depicts the escalation of congestion through the relationship between backlog, rail bottlenecks, and terminal throughput. As queues increase, rail congestion worsens, leading to higher yard congestion and reduced productivity for the port cranes. This decline in operational efficiency reduces terminal's throughput, which in turn increases queue volumes. The loop illustrates a classic positive feedback mechanism where congestion perpetuates itself through interdependent operational variables.

Surrounding this central dynamic are three balancing mechanisms. The investment-driven loop B1 captures long-term mitigation through infrastructure expansion. Rising backlog stimulates rail and terminal investments, which after a policy delay reduce congestion and stabilize throughput. The delay mark (//) signifies the temporal lag between decision-making and implementation, a critical feature of Trieste's institutional environment. The efficiency-driven loop B2 represents short-term managerial responses, such as digital scheduling and intermodal synchronization. These operational improvements enhance efficiency and throughput, counteracting the reinforcing congestion cycle. The newly introduced governance coordination loop B3 formalizes the institutional dimension: governance pressure fosters joint planning and implementation efficiency, which ultimately reduces backlog. This loop highlights the role of multi-actor coordination among the Port Authority, terminal operators, customs, and rail companies.

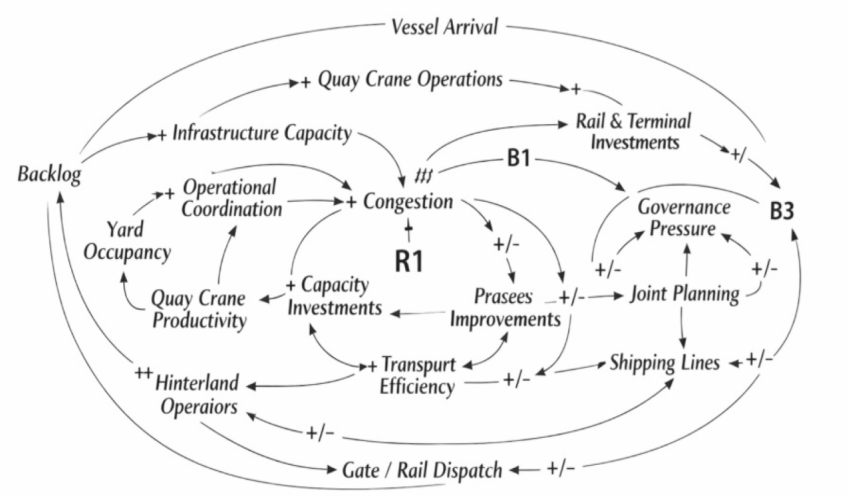


Fig. 3. Trieste-Specific Backlog Trap (Enhanced Version)

Two additional sub-structures enrich the model's realism. The modal-shift sub-loop shows adaptive behavior under rail constraints: when rail bottlenecks intensify, truck utilization rises, temporarily alleviating congestion but potentially increasing external costs such as emissions and urban traffic [33]. The hinterland feedback connects inland dispatch efficiency to effective throughput, emphasizing Trieste's dependence on hinterland evacuation capacity a distinctive characteristic of its rail-oriented logistics system [34].

Overall, the enhanced diagram advances the analytical representation of Trieste's port dynamics by integrating physical, institutional, and temporal dimensions. It distinguishes between short-term operational feedback and long-term management responses, thus offering a comprehensive framework for diagnosing systemic bottlenecks and identifying levers for sustainable improvement.

5.1 Application to the Port of Trieste

Trieste's operational and management characteristics of Trieste correspond directly to the structure presented in Fig. 3. The port's heavy reliance on rail transport means that infrastructure works, any reduction in rail access, due to a lack of rail wagons or border delays, immediately affect yard occupancy and the productivity of the port cranes. These operational dependencies correspond to the R1 reinforcing loop, where reduced outbound capacity increases congestion and delays.

Although Trieste is predominantly rail-oriented, truck-based hinterland flows also contribute to congestion dynamics. Agent-based modelling research on truck parking facilities demonstrates that parking shortages, queue spillovers, and unregulated arrival patterns can significantly increase gate congestion and extend dwell times. While these

dynamics are less dominant in Trieste compared to rail bottlenecks, they represent complementary mechanisms that can reinforce the same backlog formation processes captured in the R1 loop. Incorporating such multimodal interactions provides a more comprehensive understanding of how hinterland constraints shape terminal performance (Fig. 4) [35].

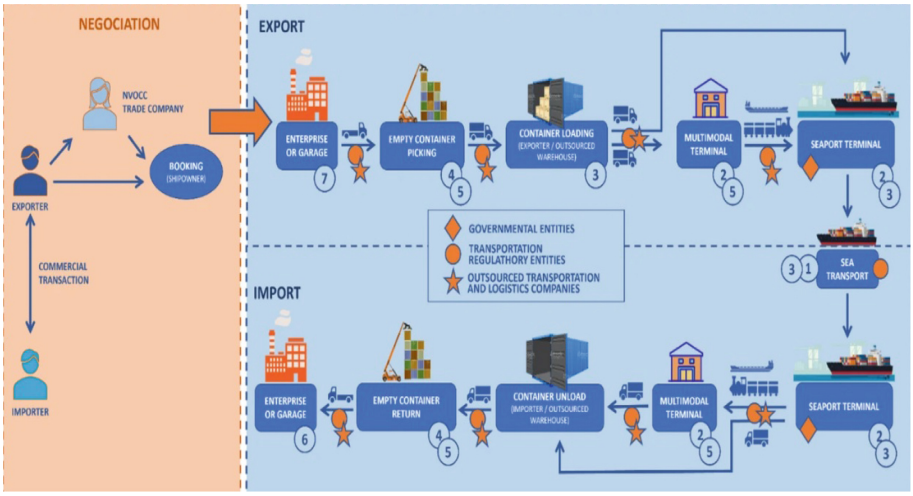


Fig. 4. Agent-based modelling of truck parking facility impacts on hinterland congestion (adapted from Reducing congestion in port hinterlands: an agent-based modelling case study of truck parking facility impacts)

The governance layer is equally significant. The port authority's strategic focus on rail development, the terminal operator's decisions on resource allocation, the scheduling practices of shipping lines, and the regulatory role of customs jointly shape the port's ability to manage peak demand. These interactions of stakeholders underpin the B1 and B2 balancing loops, where investment and efficiency improvements can alleviate congestion, but often face delays or coordination difficulties.

Considering the Trieste case within the integrated framework of Fig. 3, this section demonstrates how systemic congestion arises from the interaction of operational constraints, governance structures, and behavioral responses. The model provides a systematic framework for interpreting Trieste's performance challenges and identifying leverage points for improving resilience and efficiency.

6 Discussion

The results of this study emphasize the importance of interpreting port congestion and sustainability challenges from a systematic rather than exclusively operational perspective. Whereas traditional analysis often attributes congestion to a lack of infrastructure or to an increase in demand, the backlog trap approach shows that bottlenecks are fundamentally shaped by behavioral responses, institutional arrangements, and delayed

investments decisions. When operational pressures intensify, decision-makers prefer short-term firefighting measures over long-term capacity development, which further reinforces the condition, which leads to a buildup of deferred orders.

In the port of Trieste, this dynamic is particularly acute, as rail movements bear the main burden of cargo outbound traffic. When infrastructure works, rolling stock constraints or gaps in border coordination led to rail disruptions, their impact quickly spreads through the terminal, causing storage areas to become overloaded and reducing crane productivity. This process aligns perfectly with the R1 reinforcing loop logic, which shows how local disruptions turn into systematic congestion.

At the same time, insights from agent-based modelling studies on truck parking facilities underscore the need to consider multimodal and micro-behavioral factors. Even in rail-oriented ports such as Trieste, truck-based hinterland flows can introduce additional frictions that interact with existing bottlenecks. A shortage of parking, unregulated arrival patterns, and queue spillovers can significantly increase gate congestion and prolong dwell times, reinforcing the same delay-generation mechanisms reflected in the SD model. Including these multimodal interactions enables a more comprehensive understanding of how internal regional constraints within the shape terminal operations.

A deeper study of each specific port will require a richer empirical foundation, particularly through interviews with key actors across the operations and management chain. Such qualitative inquiry is essential for uncovering the internal and external components of congestion costs that remain hidden in aggregate indicators. Port efficiency is shaped by geographical constraints, terminal specialization, and governance models, which can generate competitive tensions even within the same port under otherwise similar conditions. These structural and institutional features raise an important question for future research: how should investment responsibilities be distributed among the different stakeholders, and in what proportion, to overcome inefficiencies and reduce the formation of accumulation? Addressing this question requires not only technical modelling but also an understanding of stakeholder incentives, power imbalances, and the economic logic guiding investment decisions within complex port ecosystems.

Taken together, the integration of System Dynamics and agent-based modelling perspectives offers a promising direction for future research. SD captures the macro-level feedback structures that drive long-term system behavior, while ABM provides granular insights into how individual actors collectively shape congestion patterns. Combining these approaches can support the design of more effective policies-such as coordinated gate appointment systems, dynamic truck parking allocation, or adaptive rail scheduling that address both structural and behavioral sources of inefficiency.

7 Conclusions

The paper presents the first results on reflections on the application of a System Dynamic Approach to ports to tackle backlog Traps, Congestion and Sustainable Transition in ports. To do that, a model is developed to analyze where and in which terms the backlog traps form.

The analysis reveals that congestion does not stem solely from a lack of physical capacity, but it arises as a systemic consequence of investments, short-term operational

priorities and coordination gap among stakeholders. These observations highlight that there is no simple technocratic or infrastructural fix for sustainable transition of ports. It is necessary to address the behavioral and governance dynamics that contribute to the increase in accumulated backlog. At the same time, a conceptual representation of the backlog trap in the port of Trieste has been presented.

The case of the Port of Trieste illustrates how structural dependencies particularly its reliance on rail-based hinterland evacuation can amplify vulnerability to disruptions. Even when rail is the dominant hinterland mode, complementary mechanisms such as truck-based flows and gate-side frictions can reinforce the same congestion loops.

Agent-based modelling studies clearly show that micro-level behavioral dynamics can accumulate and lead to a decline in system-wide efficiency. When these multi-modal observations are incorporated into the cumulative cargo burden framework, a more comprehensive picture emerges of how congestion gradually propagates from the port to inland areas and creates cascading bottlenecks.

The study suggests that further research is required to identify how congestion costs and the investments commitments are being apportioned among various stakeholders of the port. To develop effective governance and financing solutions, it is important to identify who benefits from efficiency, who bears the costs of delays and who has the resources and motivation to invest. This knowledge is fundamental to overcoming structural barriers limiting operational efficiency and decarbonization efforts.

To summarize, the results moved away from a reactive management logic and show that sustainable port performance must be addressed through systematic specialization in coordination structures including both operational integrations, on one hand being developed further into what new investment strategies should look like across stakeholder systems. Ports that fail to address backlog traps risk being locked into self-reinforcing cycle of congestion, efficiency and decarbonization delay, which significantly threatens their long-term competitiveness and environmental commitments.

Acknowledgments. PhD programme in Circular Economy at the University of Trieste, Cycle 41st, with the support of a scholarship financed by the University of Trieste.

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

Authors' Contribution Conceptualization, materials and resources all authors. Methodology, model adjustment and theoretical perspective Giuseppe Borruso.

References

1. Notteboom, T., Pallis, A., Rodrigue, J.-P.: Port Economics, Management and Policy, 1st edn. Routledge, London (2021). https://doi.org/10.4324/9780429318184
2. Smale, B.: The beginning of system dynamics
3. Sterman, J.: Business dynamics: systems thinking (2026). https://www.academia.edu/27960981/BUSINESS_DYNAMICS_SYSTEMS_THINKING_Sterman. Accessed 08 May 2026
4. Forrester, J.W.: Industrial Dynamics. Martino Fine Books, Mansfield Centre (2013)

5. Forrester, J.W.: Urban Dynamics. Cambridge, M.I.T. Press (1969) http://archive.org/details/urbandynamics0000forr. Accessed 08 May 2026
6. Forrester, J.W.: World Dynamics. Wright-Allen Press, Cambridge (1973) http://archive.org/details/worlddynamics00forr. Accessed 08 May 2026
7. Meadows, D.H.: Club of Rome: The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind. Universe Books, New York (1972) http://archive.org/details/TheLimitsToGrowth. Accessed 08 May 2026
8. Meadows, D.H.: The Limits to Growth: The 30-Year Update. Chelsea Green Publishing Company, White River Junction (2004) http://archive.org/details/limitstogrowth300000mead. Accessed 08 May 2026
9. Randers, J.: 2052: A Global Forecast for the Next Forty Years. Chelsea Green Pub, White River Junction (2012) http://archive.org/details/2052globalforeca0000rand. Accessed 08 May 2026
10. Sterman, J.: Business dynamics: systems thinking (2021). https://www.academia.edu/27960981/BUSINESS_DYNAMICS_SYSTEMS_THINKING_Sterman
11. Kacou, K., Ika, L., Munro, L.: Fifty years of capacity building: taking stock and moving research forward. Public Adm. Dev. **42**(4), 215–232 (2022). https://doi.org/10.1002/pad.1993
12. Ciobanu, S.: Edward W. Said: the policy and politics of resistance, pp. 104–111 (2017)
13. Barlas, Y.: Formal aspects of model validity and validation in system dynamics. Syst. Dyn. Rev. **12**(3), 183–210 (1996). https://doi.org/10.1002/(SICI)1099-1727(199623)12:3<183::AID-SDR103>3.0.CO;2-4
14. Ghaffarzadegan, N., Lyneis, J., Richardson, G.P.: How small system dynamics models can help the public policy process. Syst. Dyn. Rev. **27**(1), 22–44 (2011). https://doi.org/10.1002/sdr.442
15. Rahmandad, H., Sterman, J.: Heterogeneity and network structure in the dynamics of diffusion: comparing agent-based and differential equation models. Manag. Sci. **54**(5), 998–1014 (2008)
16. Eliasson, J., Mattsson, L.-G., Algers, S.: Is it time to use activity-based urban transport models? A discussion of planning needs and modelling possibilities. Ann. Reg. Sci. **39**, 767–789 (2005). https://doi.org/10.1007/s00168-005-0016-8
17. Madlener, R., Alcott, B.: Energy rebound and economic growth: a review of the main issues and research needs. Energy. **34**(3), 370–376 (2009). https://doi.org/10.1016/j.energy.2008.10.011
18. Madlener, R., Schmid, C.: Adoption and diffusion of decentralised energy conversion technologies: the success of engine co-generation in Germany. Energy Environ. **14**(5), 627–662 (2003). https://doi.org/10.1260/095830503322663384
19. Giuliano, G., O'Brien, T.: Reducing port-related truck emissions: the terminal gate appointment system at the Ports of Los Angeles and Long Beach. Transp. Res. Part D Transp. Environ. **12**(7), 460–473 (2007). https://doi.org/10.1016/j.trd.2007.06.004
20. O'Kelly, M.: The location of interacting hub facilities. Transp. Sci. **20**, 92–106 (1986). https://doi.org/10.1287/trsc.20.2.92
21. Rodrigue, J.-P., Notteboom, T.: Comparative North American and European gateway logistics: the regionalism of freight distribution. J. Transp. Geogr. **18**(4), 497–507 (2010). https://doi.org/10.1016/j.jtrangeo.2010.03.006
22. Bichou, K., Gray, R.: A logistics and supply chain management approach to port performance measurement. Marit. Policy Manag. **31**(1), 47–67 (2004)
23. Bichou, K.: Chapter 24: Review of port performance approaches and a supply chain framework to port performance benchmarking. Res. Transp. Econ. **17**, 567–598 (2006). https://doi.org/10.1016/S0739-8859(06)17024-9
24. Acciaro, M., Ghiara, H., Cusano, M.I.: Energy management in seaports: a new role for port authorities. Energy Policy. **71**, 4–12 (2014). https://doi.org/10.1016/j.enpol.2014.04.013

25. Gritsenko, D.: Regulating GHG emissions from shipping: local, global, or polycentric approach? Mar. Policy. **84**, 130–133 (2017). https://doi.org/10.1016/j.marpol.2017.07.010
26. Parola, F., Pallis, A.A., Risitano, M., Ferretti, M.: Marketing strategies of port authorities: a multi-dimensional theorisation. Transp. Res. Part A Policy Pract. **111**, 199–212 (2018). https://doi.org/10.1016/j.tra.2018.03.012
27. Woo, S.-H., Pettit, S., Beresford, A.: An assessment of the integration of seaports into supply chains using a structural equation model. Supply Chain Manag. Int. J. **18**(3), 235–252 (2013). https://doi.org/10.1108/SCM-09-2011-0264
28. Langen, P.: Governance in seaport clusters. Marit. Econ. Logist. **6**, 141–156 (2004). https://doi.org/10.1057/palgrave.mel.9100100
29. Parola, F., Satta, G., Notteboom, T., Persico, L.: Revisiting traffic forecasting by port authorities in the context of port planning and development. Marit. Econ. Logist. **23**(3), 444–494 (2021). https://doi.org/10.1057/s41278-020-00170-7
30. Bastarianto, F.F., Hancock, T.O., Choudhury, C.F., Manley, E.: Agent-based models in urban transportation: review, challenges, and opportunities. Eur. Transp. Res. Rev. **15**(1), 19 (2023). https://doi.org/10.1186/s12544-023-00590-5
31. Nguyen, T.T., et al.: Deep learning for deepfakes creation and detection: a survey. Comput. Vis. Image Underst. **223**, 103525 (2022). https://doi.org/10.1016/j.cviu.2022.103525
32. Cuong, T.N., You, S.-S., Long, L.N.B., Kim, H.-S.: Seaport resilience analysis and throughput forecast using a deep learning approach: a case study of Busan Port. Sustainability. **14**(21), 13985 (2022). https://doi.org/10.3390/su142113985
33. Lind, M., et al.: Resolving the ship backlog puzzle in the Suez Canal: predicting ship transits in capacity-constrained areas (2021)
34. Notteboom, T.E., Winkelmans, W.: Structural changes in logistics: how will port authorities face the challenge? Marit. Policy Manag. **28**(1), 71–89 (2001). https://doi.org/10.1080/03088830119197
35. Benedecti, R.C., Silva, V.M.D., da Costa, G.A.A.: Reducing congestion in port hinterlands: an agent-based modeling case study of truck parking facility impacts. J. Shipp. Trade. **11**(1), 3 (2026). https://doi.org/10.1186/s41072-026-00226-6



On the Behavior of PATE Under Class Imbalance and Non-IID Data

Alexander Bogdanov[1], Valery Khvatov[2], Yueh Chiang[1(✉)], and Nadezhda Shchegoleva[1]

[1] St. Petersburg University, St. Petersburg, Russia
st080802@spbu.ru
[2] DGT Technologies AG, Toronto, Canada

Abstract. Through systematic experimental analysis, this study investigates the behavior of PATE (Private Aggregation of Teacher Ensembles) in a federated learning (FL) context, focusing on highly class-imbalanced data and Dirichlet non-IID partitions, with the aim of exploring the applicability range of PATE. The results show that aggregation with differential privacy (DP) noise requires a sufficient number of teacher models to accumulate an adequate voting margin. However, under non-IID partitions, increasing the number of teachers dilutes minority-class samples, leading to teacher blinding and causing the margin to decrease rather than increase. When the margin falls below the noise magnitude, the aggregation outcome becomes noise-dominated. There are multi-level interactions among these factors; for tabular data with weak classification signals and severe class imbalance, the aggregated signal may fall below the noise threshold, causing a noise-dominated regime. Importantly, this behavior is attributable to the characteristics of the evaluation domain rather than to an inherent limitation of the PATE mechanism. To assist practitioners in identifying suitable configurations in advance, we propose an (S, σ/n) binning strategy as a pre-training screening indicator.

Keywords: Federated Learning · PATE · Rényi Differential Privacy · DP-SGD · Class Imbalance · Non-IID Partition

1 Introduction

Privacy regulations and cross-institutional data governance requirements limit the feasibility of centralized data collection and processing [1–3]. In this context, machine learning is shifting towards distributed, collaborative training paradigms where data remains local, such as FL [4]. FL enables participating parties to jointly train a model while retaining their raw data locally, making it a significant approach for balancing data governance and model performance. However, data localization does not eliminate privacy risks. Existing research indicates that model updates and prediction outputs can still leak information about training data, leading to attack risks such as membership inference and attribute inference [5]. Current FL methods typically collaborate through

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 327–337, 2027.
https://doi.org/10.1007/978-3-032-30539-8_25

model parameters, gradients, or updates, such as FedAvg and FedProx [4, 6]; other studies further integrate DP mechanisms, like DP-SGD, using gradient clipping and random perturbation to provide formal privacy guarantees [7]. However, under conditions of data heterogeneity and noise interference, model utility can still degrade significantly.

To reduce reliance on parameter or gradient exchange, PATE uses an ensemble of teacher models to label query samples and releases only the DP-protected aggregated labels [8]. Its utility therefore depends not only on the teacher models but also on teacher partitioning, voting structure, and the noise mechanism. Existing studies show that when teacher data shifts from approximately IID to heterogeneous partitions, teacher prediction consistency decreases, reducing the quality of aggregated labels and making ensemble decisions unstable [9].

This paper studies the applicability of PATE under federated data conditions, where class imbalance and non-IID distributions often coexist. We simulate non-IID settings using Dirichlet partitioning and control heterogeneity through the concentration parameter. Through systematic experiments, we analyze how the number of teachers, privacy budget, and data heterogeneity affect PATE performance, and identify when teacher blinding occurs. Based on this analysis, we propose an (S, σ/n) binning strategy as a practical pre-training indicator for screening PATE applicability and unstable configurations.

2 Foundational Analysis

2.1 Basic Definition of DP

Two datasets D and $D^{'}$ differing in at most one record are called neighboring datasets, denoted as $D \sim D^{'}$. A randomized mechanism M satisfies ϵ-DP if for any set of measurable events S, the following condition holds [10].

$$\Pr[M(D) \in S] \leq e^{\epsilon} \Pr\left[M\left(D'\right) \in S\right] \tag{1}$$

To calibrate the noise, the maximum impact of a single record on the query output must be quantified. For a vector-valued query function f, its L_1 sensitivity is defined as follows.

$$\Delta_1(f) = \max_{D \sim D'} \|f(D) - f\left(D'\right)\|_1 \tag{2}$$

2.2 Laplace and Gaussian Mechanisms

Under the L_1 sensitivity setting, the query result can be released using the Laplace mechanism:

$$M(D) = f(D) + Y, Y_i \sim Lap\left(\frac{\Delta_1(f)}{\epsilon}\right) \tag{3}$$

When a small failure probability δ is permissible, (ϵ, δ)-DP can be employed:

$$\Pr[M(D) \in S] \leq e^{\epsilon} \Pr\left[M\left(D'\right) \in S\right] + \delta \tag{4}$$

Correspondingly, the L_2 sensitivity of the query is defined as follows.

$$\Delta_2(f) = \max_{D \sim D'} \|f(D) - f\left(D'\right)\|_2 \tag{5}$$

If Gaussian noise is added to the output, the mechanism can be written as:

$$M(D) = f(D) + Z, Z \sim N\left(0, \sigma^2 I\right) \tag{6}$$

Compared to the Laplace mechanism, the Gaussian mechanism is more suitable for analyzing vector-valued vote statistics and multiple sequential queries.

2.3 Composition of Repeated Queries and RDP Accounting

In scenarios involving multiple sequential queries, the privacy cost accumulates with the number of queries. For (ϵ, δ)-DP, if the Gaussian mechanism is applied k times to the same dataset, their sequential composition results in $(k\epsilon, k\delta)$, where ϵ is the privacy parameter for a single query. Under the RDP framework [11], the privacy loss of the Gaussian mechanism can be expressed as:

$$\epsilon_{RDP}(\alpha) = \frac{\alpha \Delta^2}{2\sigma^2} \tag{7}$$

If the total number of queries is k, the cumulative cost is

$$\epsilon_{RDP}^{(k)}(\alpha) = k \cdot \frac{\alpha \Delta^2}{2\sigma^2} \tag{8}$$

This can then be converted back to (ϵ, δ)-DP using:

$$\epsilon = \epsilon_{RDP}^{(k)}(\alpha) + \frac{\log(1/\delta)}{\alpha - 1} \tag{9}$$

Since ϵ depends on α, it can be minimized for $\alpha > 1$ to obtain the corresponding privacy loss estimate. This method supports a larger number of queries under the same privacy budget.

2.4 PATE Voting Query and Sensitivity

In PATE, n teachers classify an unlabeled sample x. The vote count for class c can be represented as [8]:

$$V_c(x) = \sum\nolimits_{k=1}^{n} 1\left[f_k(x) = c\right] \tag{10}$$

This paper uses the vote histogram consisting of the vote counts for each class as the query output and adds Gaussian noise before releasing the final label:

$$M(x) = \operatorname*{argmax}_c (V_c(x) + Z_c), Z_c \sim N\left(0, \sigma^2\right) \tag{11}$$

If a single teacher's prediction for sample x changes from class i to class j, the count for one class decreases by one and the count for the other increases by one, resulting in a sensitivity $\Delta_2 = \sqrt{2}$.

2.5 Voting Margin Analysis

Consider a binary classification problem. Let the total number of training samples be N and the proportion of the minority class be π. Then the total number of minority class samples is $N_1 = \lfloor \pi N \rfloor$. We are concerned with whether the minority class samples become excessively diluted due to partitioning, thus failing to form a reliable aggregation capable of resisting noise.

Under class-wise Dirichlet partitioning, the allocation of minority class samples is a two-step process. First, the data proportion for each teacher is sampled from a Dirichlet distribution. Second, the N_1 minority class samples are assigned to each teacher according to this proportion. Therefore, the number of minority class samples C_k assigned to a single teacher follows a Beta-Binomial distribution:

$$P(C_k = j) = \binom{N_1}{j} \frac{B(j+\beta, N_1 - j + (n-1)\beta)}{B(\beta, (n-1)\beta)} \tag{12}$$

To simplify the analysis, taking logistic regression as an example, we use the Events Per Variable (EPV) criterion as a condition to determine whether a teacher can effectively learn the minority class. Each teacher needs at least $10d$ minority class samples for effective learning, where d is the number of predictor variables. In this paper, we set the threshold $k_{EPV} = 30$, which corresponds to the conservative engineering setting $d = 3$ effective predictors. For datasets with higher feature dimensionality, the threshold scales proportionally: at $d = 10$, $k_{EPV} \approx 100$; at $d = 20$, $k_{EPV} \approx 200$. In all experiments reported here, d is set to match the feature count of the controlled dataset generated by sklearn. make classification, so the threshold is re-evaluated per dataset rather than applied globally. Define $L_k = 1_{C_k \geq 30}$ as an indicator of whether teacher k is qualified, and $L = \sum_{k=1}^{n} L_k$ as the total number of qualified teachers.

Under class-wise Dirichlet partitioning, the number of minority class samples C_k assigned to a single teacher follows a Beta-Binomial distribution, so the probability that any given teacher is qualified is (Fig. 1):

$$p_{30}(n) = P(C_k \geq 30) = 1 - \sum_{j=0}^{29} \binom{N_1}{j} \frac{B(j+\beta, N_1 - j + (n-1)\beta)}{B(\beta, (n-1)\beta)} \tag{13}$$

The expected number of qualified teachers is $E[L] = np_{30}(n)$.

Adopting a most optimistic voting assumption: qualified teachers correctly vote for the minority class, while unqualified teachers vote for the majority class. Then, given L, the vote count for the minority class is $V_1 = L$, and the vote count for the majority class is $V_0 = n - L$. The voting margin is defined as $m = V_1 - V_0 = 2L - n$. Since L is a random variable, we define its expected voting margin as $E[m] = 2E[L] - n$. Let $p_{30}(n) = P(C_k \geq 30)$, then $E[L] = np_{30}(n)$, thus $E[m] = n(2p_{30}(n) - 1)$.

During PATE aggregation, independent Gaussian noise $Z_1, Z_0 \sim N(0, \sigma^2)$ is added to the vote counts for both classes. The condition for correct decision-making is $V_1 + Z_1 > V_0 + Z_0$, i.e., $Z_0 - Z_1 < m$. Since $Z_0 - Z_1 \sim N(0, 2\sigma^2)$, the accuracy given m is

$$P(correct \vee m) = \Phi\left(\frac{m}{\sqrt{2}\sigma}\right) \tag{14}$$

Requiring accuracy at least $1-\delta$ necessitates $m \geq \sqrt{2}\sigma\Phi^{-1}(1-\delta)$. Define $\tau_{\sigma,\delta} = \sqrt{2}\sigma\Phi^{-1}(1-\delta)$. Even without considering data partitioning, the maximum margin $m \leq n$ requires $n \geq \tau_{\sigma,\delta}$. Considering data partitioning, we need

$$n(2p_{30}(n)-1) \geq \tau_{\sigma,\delta} \tag{15}$$

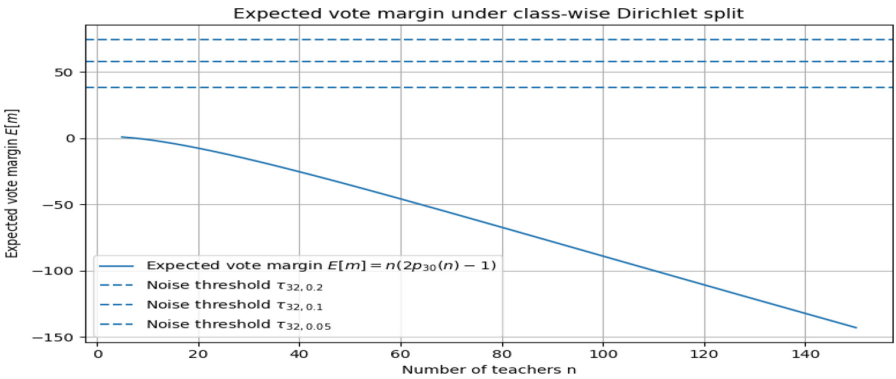


Fig. 1. Voting Margin vs. Number of Teachers

This figure shows the change in the expected voting margin $E[m] = n(2p_{30}(n) - 1)$ as the number of teachers n increases under class-wise Dirichlet partitioning ($\beta = 0.3$). It can be observed that as n increases, $E[m]$ monotonically decreases and quickly becomes negative, indicating that most teachers tend to vote for the majority class due to insufficient minority class samples ($C_k < k_{epv}$), causing the overall vote to lean towards the incorrect class.

Additionally, the figure uses dashed lines to mark the noise thresholds $\tau_{32,\delta}$ of the Gaussian mechanism for different δ values. Across the entire range of n examined, $E[m]$ is far below $\tau_{32,\delta}$, and is even negative in most regions, demonstrating that under these conditions of data imbalance ($\pi = 0.067$) and partitioning, the aggregated signal cannot form an effective margin sufficient to resist noise.

This result indicates that when the number of teachers increases, the dilution effect on the minority class caused by Dirichlet partitioning dominates the system's behavior, driving the PATE aggregation into a noise-dominated regime in practice and making reliable decision-making impossible. It is worth noting that, for tabular data, when the number of predictor variables is larger or the number of classes in the classification problem is higher, the model requires more minority-class samples to learn a stable decision boundary; however, the required sample size is difficult to estimate in an engineering manner. Therefore, this paper subsequently proposes the (S, σ/n) binning strategy as a practical screening indicator for evaluating the applicability of PATE.

3 Experimental Evaluation

3.1 Experimental Setup

The first part of the experiment uses one public financial benchmark dataset: GiveMeSomeCredit (sample size N = 10,000, feature dimension d = 10, minority-class proportion π =0.067) [12]. The dataset achieved a centralized no-privacy baseline AUC of 0.784 using an 80/20 stratified train-test split. This part illustrates that, when the sample size is limited, classes are imbalanced, and the privacy budget is low, uncalibrated parameter settings may drive PATE into a performance degradation regime. The second part further includes German Credit (N = 1,000, d = 20, π =0.30) [13], Bank Marketing (N = 41,188, d = 20, π =approximately 0.113) [14], and the CDC Diabetes Dataset (N = 253,680, d = 21, π =0.139) [15], together with the previous dataset as auxiliary experiments for illustrating mitigation strategies under engineering settings. Their centralized no-privacy baseline AUCs under the same 80/20 stratified train-test split were 0.779, 0.951, and 0.823, respectively.

To simulate a non-IID environment, this paper uses the Dirichlet distribution for data partitioning, with the concentration parameter fixed at $\alpha = 0.3$. All methods are compared under the same data partitioning, with privacy settings following (ε, δ)-DP, where $\delta = 10^{-5}$.

3.2 Baseline Comparison with DP-SGD

The first part of the experiment uses the GiveMeSomeCredit dataset and logistic regression teacher models, with three teacher-count settings (10, 30, and 50), query counts of 20, 30, and 30, respectively, and privacy budgets $\varepsilon \in \{1, 2, 5, 10\}$.

With fixed data volume, the left panel of Fig. 2 compares the logistic regression PATE results with different numbers of teachers on GiveMeSomeCredit. DP-SGD maintains an AUC between approximately 0.64 and 0.69 across privacy budgets. In contrast, the AUC for the three PATE-LR configurations is generally lower, ranging approximately from 0.31 to 0.55. With 10 teachers, the AUC is approximately 0.55 at $\varepsilon = 2$; with 30 teachers, it ranges approximately from 0.40 to 0.44; with 50 teachers, it ranges approximately from 0.35 to 0.45. Overall, under this setting, increasing the number of teachers does not show a consistent improvement in performance.

On the other hand, the right panel of Fig. 2 presents the proportion of blind teacher models. Under Dirichlet partitioning, as the number of teachers increases, the blind-teacher proportion rises from 0.07 with 10 teachers to 0.28 and 0.27 with 30 and 50 teachers, respectively, consistent with the expected dilution of minority-class samples.

3.3 Mitigation Strategy under Engineering Conditions

The preceding results indicate that, under weak class signals and non-IID partitioning, PATE may enter a performance degradation regime due to teacher blinding and insufficient voting margins. We therefore evaluate whether the average cosine similarity S between teacher data distributions and σ/n can serve as practical binning indicators for screening PATE applicability. Here, S describes the degree of similarity among teacher

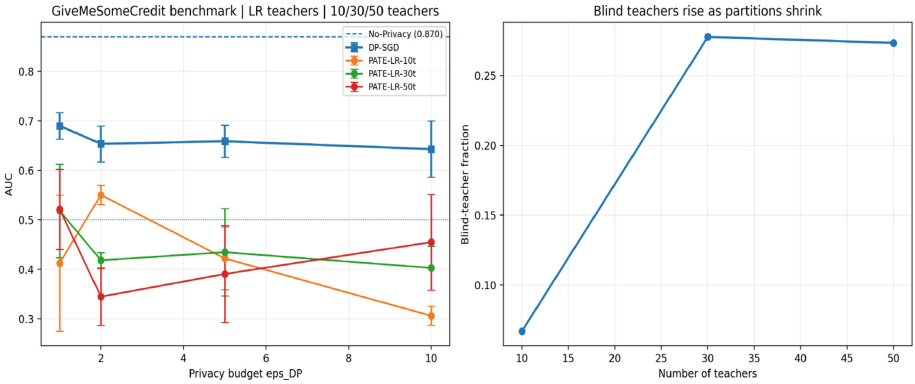


Fig. 2. Teacher Count vs. Performance and Blind Ratio

data distributions, while σ/n reflects the ratio of noise intensity to the number of teachers. For each real dataset, we generate a controlled dataset using sklearn.make_classification, matched in minority-class proportion π, sample size N, and feature dimension d, with class_sep adjusted according to the original-data baseline AUC. German Credit uses a query pool of 200 samples, while the other datasets use 2,000 samples. Query samples are clustered before selection to avoid all-negative query batches. Experiments are conducted under a fixed privacy budget with $\alpha \in \{0.1, 0.3, \infty\}$, $n \in \{20, 40, 60, 80, 100, 125, 150, 175, 200\}$, $Q \in \{20, 40, 60, 80, 100\}$, and 16 random seeds. A configuration is treated as a failure if any student model has an AUC below 0.7, and failure ratios are computed over (S, σ/n) bins to construct the heatmaps.

As shown in the figures, the heatmaps for all datasets exhibit continuous regional structures rather than completely random distributions, indicating that the (S, σ/n) binning strategy has a certain ability to distinguish regions where PATE performance degrades. The learnable regions usually appear as contiguous areas rather than scattered points, suggesting that stable PATE performance is concentrated mainly within specific parameter ranges. Although the S-bin distributions of the real and controlled data differ and the grid cells do not align perfectly, the heatmap structures remain clearly similar.

The analysis of the GiveMeSomeCredit dataset in Fig. 3 echoes the preceding results: for a low-data-volume dataset, there is almost no learnable region. German Credit in Fig. 4 is an even smaller dataset, and the conclusion is the same. In contrast, Bank Marketing in Fig. 5 and CDC Diabetes in Figure 6 retain a wider range of low-failure-rate regions as the data volume increases and the degree of imbalance decreases. Although the grid cells do not align perfectly, this binning approach still presents regional trends close to those of the real data, indicating that even under simplified distributional conditions, controlled data can provide an approximate evaluation of the learnable and degraded regions of PATE.

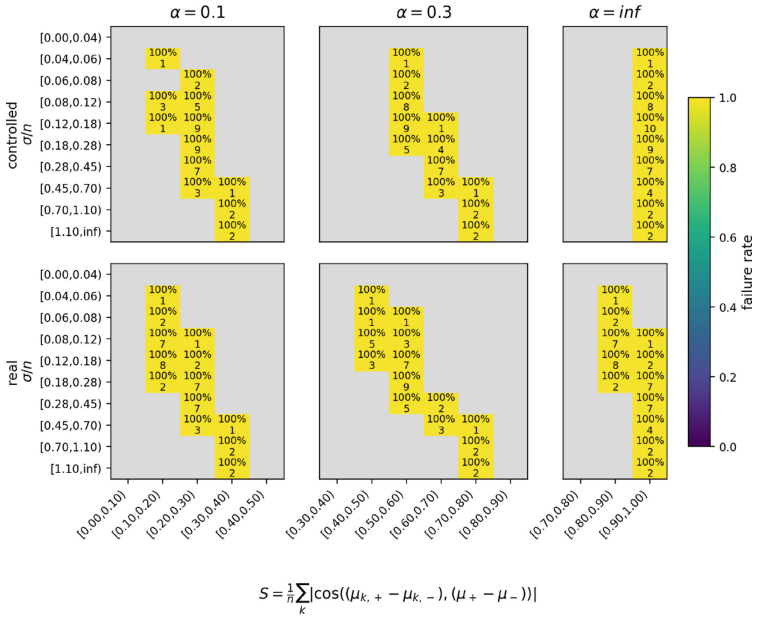


Fig. 3. Heatmap for GiveMeSomeCredit

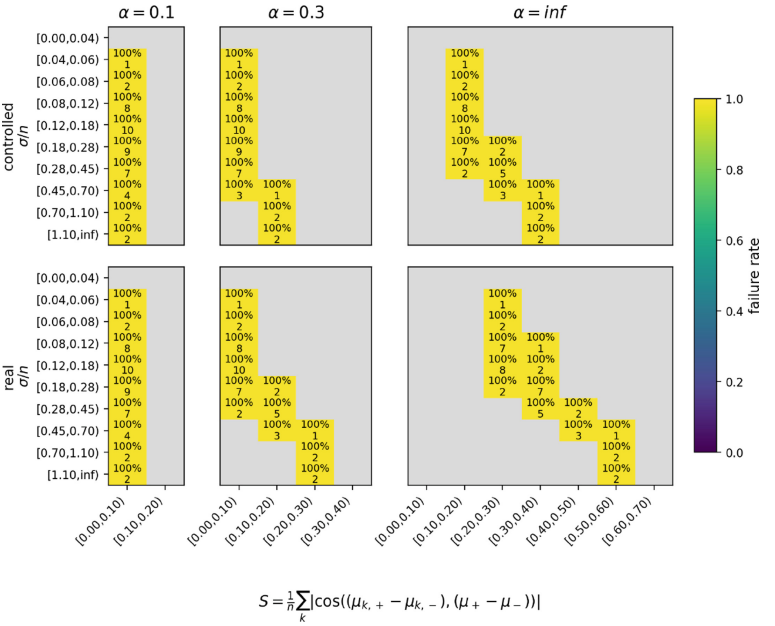


Fig. 4. Heatmap for German Credit

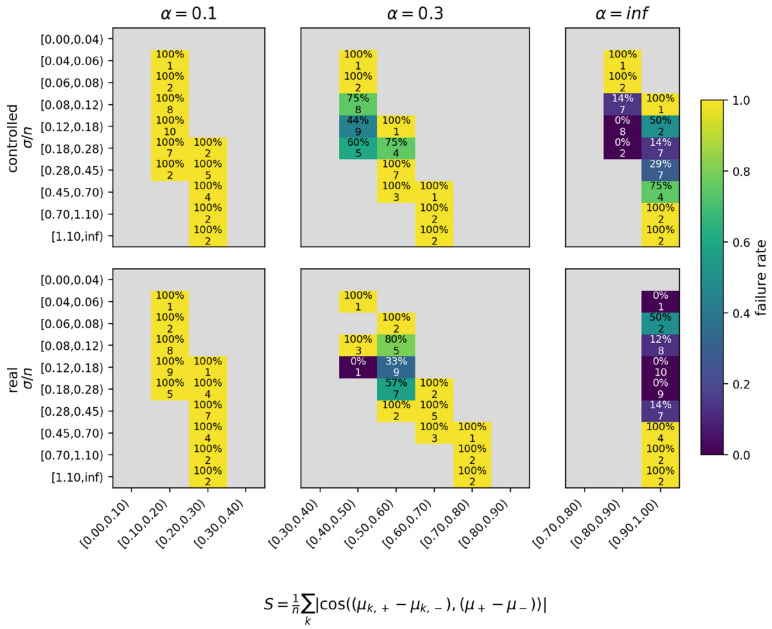


Fig. 5. Heatmap for Bank Marketing

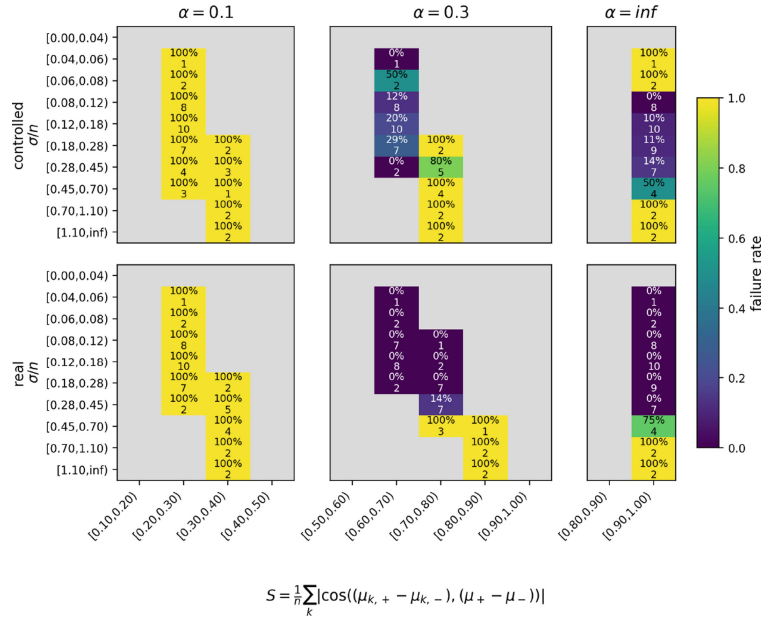


Fig. 6. Heatmap for CDC Diabetes

4 Conclusion

In summary, this study demonstrates that the performance of PATE in FL scenarios is closely related to data distribution, class structure, and signal redundancy. When minority-class samples are diluted due to Dirichlet partitioning, the teacher voting margin decreases as the number of teachers increases, causing the aggregation process to enter a noise-dominated region. On tabular data with weak class signals, this phenomenon may lead to noticeable performance degradation.

To assist practitioners in determining whether a specific dataset is suitable for PATE, this paper proposes an (S, σ/n)-based binning strategy as a pre-training applicability screening indicator. We find that the failure-rate heatmaps generated by this binning approach exhibit similar regional structures, indicating that the strategy can effectively distinguish the learnable regions of PATE from regions of performance degradation.

However, the results of this paper do not imply that PATE is unsuitable for high-dimensional data. On the contrary, the original PATE work was mainly built on datasets such as MNIST, which contain high-dimensional redundant signals, and subsequent studies such as SeqPATE [16] are likewise based on high-dimensional semantic structures in language models. In such data, even in the presence of DP noise and distributed partitioning, teacher models may still maintain relatively high consistency and stable voting margins.

Therefore, this paper argues that the performance degradation observed in this study mainly results from voting-margin decay jointly caused by weak classification signals, class imbalance, and non-IID partitioning, rather than from a failure of the PATE method itself. Based on this observation, future work will further explore federated text-learning scenarios with stronger semantic redundancy and classification signals, in order to analyze the applicability and stability of PATE under different data conditions.

References

1. European Parliament and Council of the European Union: Regulation (EU) 2016/679 (General Data Protection Regulation). Off. J. Eur. Union (2016). https://eur-lex.europa.eu/eli/reg/2016/679/oj
2. Standing Committee of the National People's Congress: Personal Information Protection Law of the People's Republic of China (2021). https://www.cac.gov.cn/2021-08/20/c_1631050028355286.htm
3. European Parliament and Council of the European Union: Regulation (EU) 2024/1689 (Artificial Intelligence Act). Off. J. Eur. Union (2024). https://eur-lex.europa.eu/eli/reg/2024/1689/oj
4. McMahan, H.B., Moore, E., Ramage, D., Hampson, S., Aguera y Arcas, B.: Communication-efficient learning of deep networks from decentralized data. In: Proceedings of 20th International Conference on Artificial Intelligence and Statistics (AISTATS), pp. 1273–1282 (2017)
5. Shirvani, G., Ghasemshirazi, S., Beigzadeh, B.: Federated learning: attacks, defenses, opportunities, and challenges. In: Proceedings of the 11th International Symposium on Telecommunications (IST), pp. 470–479 (2024)
6. Li, T., Sahu, A.K., Zaheer, M., Sanjabi, M., Talwalkar, A., Smith, V.: Federated optimization in heterogeneous networks. In: Proceedings of Machine Learning and Systems (MLSys), pp. 429–450 (2020)

7. Abadi, M., et al.: Deep learning with differential privacy. In: Proceedings of the ACM Conference on Computer and Communications Security (CCS), pp. 308–318 (2016)
8. Papernot, N., Abadi, M., Erlingsson, Ú., Goodfellow, I.J., Talwar, K.: Semi-supervised knowledge transfer for deep learning from private training data. In: International Conference on Learning Representations (ICLR) (2017)
9. Dodwadmath, A., Stich, S.U.: Preserving privacy with PATE for heterogeneous data. In: NeurIPS 2022 Workshop on Distribution Shifts (DistShift) (2022) https://openreview.net/forum?id=NvlUqen8Cya
10. Dwork, C., Roth, A.: The algorithmic foundations of differential privacy. Found. Trends Theor. Comput. Sci. **9**(3–4), 211–407 (2014)
11. Mironov, I.: Rényi differential privacy. In: Proceedings of the IEEE Computer Security Foundations Symposium (CSF), pp. 263–275 (2017)
12. Credit Fusion, Cukierski, W.: Give me some credit [Dataset]. Kaggle Competition (2011). https://kaggle.com/competitions/GiveMeSomeCredit
13. Hofmann, H.: Statlog (German Credit Data) [Dataset]. UCI Machine Learning Repository (1994). https://doi.org/10.24432/C5NC77
14. Moro, S., Rita, P., Cortez, P.: Bank Marketing [Dataset]. UCI Machine Learning Repository (2014). https://doi.org/10.24432/C5K306
15. Centers for Disease Control and Prevention (CDC): Behavioral Risk Factor Surveillance System 2015 [Dataset] (2015). https://www.cdc.gov/brfss/annual_data/annual_2015.html
16. Tian, Z., Zhao, Y., Huang, Z., Wang, Y.X., Zhang, N.L., He, H.: SeqPATE: differentially private text generation via knowledge distillation. In: Advances in Neural Information Processing Systems (NeurIPS), vol. 35, pp. 11117–11130 (2022)

Traffic Flow Analysis of Closely Spaced Roundabouts and Signalized Intersections

Mirna Klobučar[1], Aleksandra Deluka-Tibljaš[1](✉), Sanja Šurdonja[1], and Irena Ištoka Otković[2]

[1] Faculty of Civil Engineering, University of Rijeka, Rijeka, Croatia
aleksandra.deluka@uniri.hr
[2] Faculty of Civil Engineering and Architecture Osijek, Josip Juraj Strossmayer University of Osijek, Osijek, Croatia

Abstract. Heterogeneous traffic flow conditions in urban areas often result in reduced traffic efficiency, safety issues, and increased environmental impacts. Roundabouts are often seen as solutions to all of these issues; however, when located in close proximity to signalized intersections, their operational advantages may be significantly reduced. The main issue associated with the close spacing of roundabouts and signalized intersections arises from interrupted vehicle movements at signalized intersections, in contrast to the continuous flow characteristic of roundabouts. This paper presents part of a broader research focused on analyzing operational performance and optimizing the distance between roundabouts and signalized intersections. The presented study is based on a real-world traffic network in a small, highly touristic town in Croatia (Poreč). A VISSIM traffic microsimulation model was developed and calibrated, using travel time data, for a network segment consisting of a roundabout and a signalized intersection spaced 125 m apart. Traffic performance was evaluated using selected traffic indicators. The analysis included the existing traffic configuration and a hypothetical scenario in which the roundabout is functioning independently. Both measured traffic demand and a theoretically defined maximum traffic load (ADT 20,000) were considered under identical geometric and location conditions. The results provide insights into the impact of closely spaced intersections on traffic performance and support the need for careful consideration of intersection spacing in urban traffic planning.

Keywords: roundabout · intersection interaction · traffic microsimulations · urban network · signalized intersection

1 Introduction

Roundabouts are often seen as solutions for both traffic flow and traffic safety challenges in urban road networks. However, when located in close proximity to signalized intersections, their operational advantages may be significantly reduced. The main issue associated with the close spacing of roundabouts and signalized intersections arises from interrupted vehicle movements at signalized intersections, in contrast to the continuous

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 338–350, 2027.
https://doi.org/10.1007/978-3-032-30539-8_26

flow characteristic of roundabouts. Although this problem has been recognized by the professional and scientific community, it has not been researched in detail, and there are no clear recommendations regarding the conditions under which roundabouts and signalized intersections can be combined in the same corridor.

Technical guidelines and regulations in Europe and worldwide predominantly address isolated roundabouts, focusing on geometric design elements and performance comparisons with conventional at-grade intersections. A broader international comparison of roundabout design practices was presented by Nikiforiadis et al. [1], who reviewed roundabout guidelines from the UK, France, Germany, the Netherlands, the USA, and Australia. Their study showed that, although similar geometric design elements are commonly considered, the detailed design principles and recommended values differ among countries. They also emphasized that roundabout design involves balancing capacity and safety, since increasing entry, exit, or circulatory lanes may improve capacity but can also increase vehicle speeds and reduce safety. These differences indicate that roundabout performance cannot be evaluated solely on the basis of isolated geometric design characteristics, but must also be considered in relation to local traffic demand, operational conditions, and the wider network context.

A similar analysis of the guidelines from selected EU countries and the USA, from the aspect of the functioning of the roundabout adjacent to the signalized intersection, done in this research, shows that the recommendations generally conclude that such a combination is not advisable, though they often lack more detailed explanations [2–5]. In German [2], Norwegian [5], and UK [5] guidelines the use of a double (or large) roundabout is suggested when the center-to-center spacing is less than 100 meters, with a minimum allowable spacing of 40 meters in the double-roundabout configuration. The US guidelines [6] define a distance of less than 150 m as critical between the roundabout and signalized intersection, and in such cases, a detailed corridor-level microsimulation analysis is recommended.

Research on this topic in the available scientific literature is still limited and relatively recent.

In 2025 [7] a study was published on two corridors in different countries (Italy, France) under similar conditions in which a series of urban traffic signalized intersections were partially or completely replaced by roundabouts. The results showed that corridors composed entirely of roundabouts do not necessarily provide the best balance between traffic performance and safety. In Pisa, the most efficient configuration included four roundabouts and one signalized intersection, while the safest included three roundabouts and two signalized intersections. In Avignon, the best results were achieved with roundabouts combined with strategically placed signalized intersections. Consequently, the study highlights that optimal corridor performance is achieved not through uniformity, but through carefully balanced combinations tailored to local traffic and safety conditions.

Previous research by this group of authors based on a theoretical traffic microsimulation model of the interaction of roundabouts and traffic signalized intersections [8] in different traffic and distance conditions, enabled the definition of parameters that significantly affect the quality of traffic flows when a roundabout is in close proximity to signalized intersection. By varying the distance between intersections (40-120 m), the

volume and distribution of traffic, it was found that all of the above factors affect the quality of traffic flows. The impact of intersection spacing was particularly pronounced when it came to smaller distances (40-60 m), while at a distance of 100 m the intersection combination had almost the same traffic performance as the case of isolated roundabouts under the same traffic conditions. This research motivated the investigation of the interaction of roundabouts and traffic lights based on real traffic conditions presented in this paper.

This study is based on a real traffic network in Poreč, a small tourist town in Croatia. A VISSIM microsimulation model was developed and calibrated using travel time data for a road segment that includes a roundabout and a signalized intersection spaced 125 m apart. Traffic performance was assessed using selected indicators under both the existing configuration and a hypothetical scenario in which the roundabout operated independently. The aim was to evaluate the impact of the signalized intersection on roundabout performance. Both measured traffic demand and a theoretical maximum traffic load (ADT 20,000) were analyzed under the same geometric and locational conditions. The results highlight the effects of closely spaced intersections on traffic performance using real-world data.

The structure of the paper is organized as follows. Sect. 2 details the materials and methods, including field measurements, the development and calibration of the VISSIM microsimulation model, and the definition of the analyzed scenarios. Section 3 presents and compares the simulation results for both the existing roundabout-signalized intersection configuration and the isolated roundabout scenario across varying traffic demand levels. Finally, Section 4 summarizes the key findings, discusses their implications for urban traffic planning, and suggests directions for future research.

2 Materials and Methods

The methodological framework of this study is presented in Fig. 1. Field measurements were conducted in Poreč, Croatia, to obtain design and traffic parameters for the analyzed network segment, which consists of a roundabout and a signalized intersection. Based on the field data, a traffic microsimulation model was developed in PTV VISSIM. The model was calibrated against observed travel time data to achieve a reliable representation of existing traffic conditions. The analysis included the existing traffic configuration, which reflects the interaction between the roundabout and the signalized intersection, as well as a hypothetical scenario in which the roundabout operated as an isolated intersection. For each scenario, two traffic volume levels were considered: the measured real traffic volume and a theoretically defined maximum traffic load of 20,000 vehicles per day (ADT 20,000). Traffic performance was then evaluated using selected traffic indicators.

2.1 Field Measurements

The study area is situated in the city of Poreč, Croatia, within a predominantly residential area adjacent to the historic town center, a shopping plaza, and the main municipal parking facility (Fig. 2).

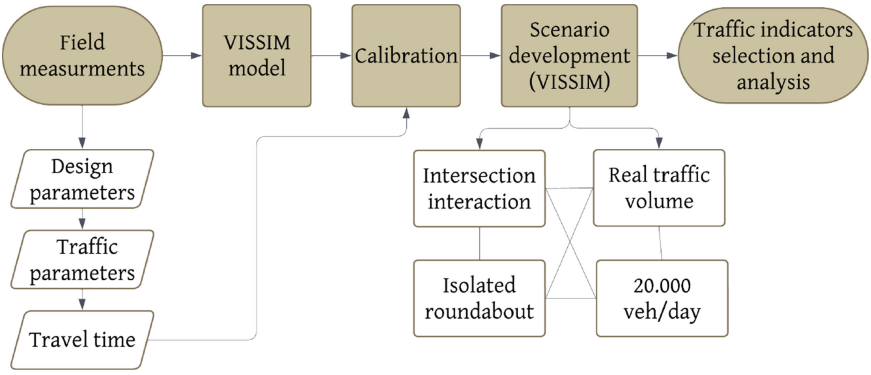


Fig. 1. Research methodology

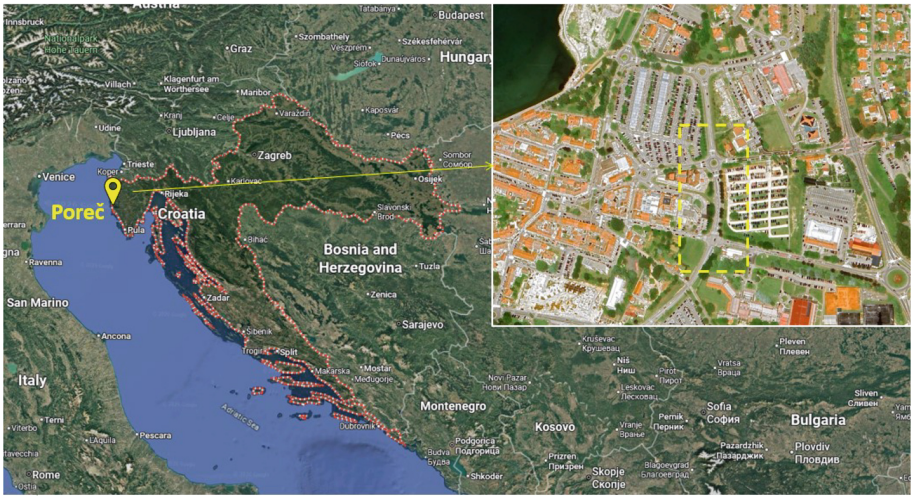


Fig. 2. Geographical location of the case study area (Google Maps)

The analyzed network segment comprises a roundabout and a signalized intersection located 125 m apart, as shown in Fig. 3.

The measurements were conducted in July and as Poreč is a well-established summer tourist destination, a considerable proportion of traffic at these intersections was generated by tourists, leading to increased variability in driver behavior.

The analyzed roundabout has an outer radius of 18 meters and, according to the Croatian Guidelines for Roundabout Design [9], is classified as a medium-sized urban roundabout. Under these guidelines, the recommended maximum capacity (ADT) for a roundabout of this size is 20,000 vehicles per day. Figure 4 shows the geometric elements of the analyzed roundabout. The adjacent signalized intersection is a four-leg intersection operating with a two-phase signal timing plan. The cycle length is 70 s, with



Fig. 3. Analyzed network segment (Google Maps)

the green interval of the first phase lasting 40 s, followed by 2 s of red-amber, 25 s of red, and 2 s of amber.

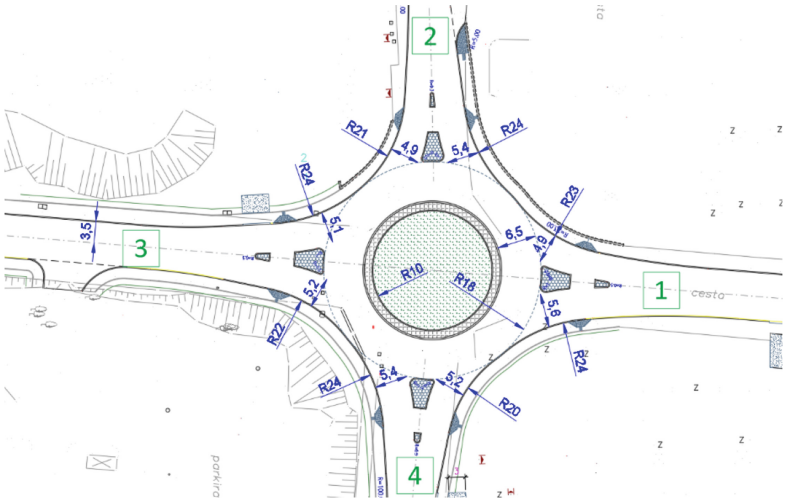


Fig. 4. Design elements of the roundabout

In addition to design elements of the roundabout, field measurements were conducted to collect data on traffic volume and traffic structure using equipment from the Transportation Laboratory of the Faculty of Civil Engineering, University of Rijeka. Traffic counters Datacollect SRD were used to continuously record traffic flow on every approach (in both directions) over a 24-hour period on Tuesday, July 16 2024. The results are shown in Table 1.

Table 1. Traffic data

Approach	Traffic structure			Daily traffic volume	Peak traffic volume (*veh/h)*	
	car	*truck*	*HGV*	*veh/day*	07:00-08:00	15:00-16:00
1	97.3%	2.7%	0.0%	2208	132	141
2	97.7%	2.3%	0.0%	3139	193	275
3	98.1%	1.7%	0.1%	5063	312	357
4	98.8%	1.2%	0.0%	3542	202	218

A daily traffic volume of 13,952 vehicles was recorded, with 991 vehicles observed during the afternoon peak hour, accounting for 7.1% of the daily total. Passenger cars account for the largest share of vehicles, while the proportion of heavy goods vehicles (HGVs) is almost negligible.

Travel time was measured during the afternoon peak hour (15:00 – 16:00) to evaluate the travel times between designated control points, specifically from Point 1 to Point 2, as shown in Fig. 2, along a 300-meter route. The data collection was carried out using the GeoTracker GPS application, which provides high-resolution measurements for both route length and travel duration. A passenger vehicle was driven by three different drivers to collect the data. Travel times were recorded from 15 independent vehicle passes through both the roundabout and the adjacent signalized intersection. Descriptive statistics for the measured travel times are summarized in Table 2.

Table 2. Descriptive statistics of measured travel time

	Nbr. of observations	Minimum	Maximum	Median	Mean	Standard deviation
Travel time (s)	15	47.00	107.00	54.00	60.33	15.32

2.2 Development and Calibration of the VISSIM Model for the Existing Traffic Configuration

Based on field measurements conducted in the city of Poreč, a microsimulation model of a selected segment of the traffic network was developed using VISSIM (Fig. 5).



Fig. 5. VISSIM model of the analyzed traffic segment

PTV VISSIM is a microscopic, multimodal traffic simulation software that models the movements and interactions of individual vehicles, pedestrians, cyclists, and public transport with great detail.

The analyzed network segment was modeled in PTV VISSIM using links and connectors to reproduce the geometry of the roundabout, the connecting road section, and the adjacent signalized intersection. Vehicle inputs were assigned according to the measured approach volumes, while the observed traffic composition was used to define the vehicle classes. The existing two-phase signal-control plan was implemented using the measured cycle length and phase durations. Priority rules and conflict areas were adjusted to reproduce vehicle interactions at the roundabout entries and along the short section between the roundabout and the signalized intersection.

To ensure data accuracy, the VISSIM model was calibrated using travel time data. Travel time was chosen because it encompasses key aspects of driver behavior, including desired speeds, reaction times, vehicle interactions, and queue formation, thereby providing a comprehensive overview of traffic flow dynamics. Moreover, this metric is relatively straightforward and offers precise measurements. Existing research supports the use of travel time as an effective calibration criterion across diverse traffic conditions [10–12].

The model calibration regarding travel time was conducted over a 300 m route connecting designated points 1 and 2 (Fig. 3). This route included traversing a roundabout and a signalized intersection. The calibration process involved comparing simulated travel times with actual travel times recorded during field measurements (Sect. 2.1). The calibration criterion [10] was defined by Eq. (1):

$$\left| \frac{T_{MOD} - T_{MEAS}}{T_{MEAS}} \right| \leq 5\% \tag{1}$$

where T_{MOD} presents the mean model-predicted travel time between reference points, and T_{MEAS} indicates the mean measured travel time between the same points. The model is considered calibrated when the above condition is satisfied.

For calibration, the following parameters were adjusted:

- Random seed increment (default=2, calibrated=6)

- Average standstill distance (m) (default=2, calibrated=2.8)
- Additive part of the desired safety distance (m) (default=2, calibrated=2.7)
- Multiplicative part of the desired safety distance (m) (default=2, calibrated=3.5)

The random seed increment determines the step size for the random number generator, resulting in different stochastic values within the model. The other three parameters pertain to driver behavior as outlined by the Wiedemann 74 car-following model, which is particularly well-suited for urban traffic networks.

The results of the travel time comparison are shown in Table 3.

Table 3. Comparison of travel times – VISSIM models and field measurements

	T_{MEAS} (s)	T_{MOD} (s)	Relative deviation
Route (300 m)	60.33	59.30	1,71 %

2.3 Development of Alternative Scenarios Using the Traffic Microsimulations

Following model calibration, additional scenarios were developed to examine the effects of traffic demand and of the spacing between the roundabout and the signalized intersection on key traffic performance measures.

A total of four scenarios were analyzed, as shown in Fig. 1:

- Scenario 1 (RB-SI) – Existing site configuration with the current traffic volume of 13,952 veh/day
- Scenario 2 (RB-SI_20k) – Existing site configuration with increased roundabout traffic volume of 20,000 veh/day
- Scenario 3 (Isolated RB) – Isolated roundabout scenario with the current traffic volume
- Scenario 4 (Isolated RB_20k) – Isolated roundabout scenario with increased traffic volume of 20,000 veh/day

In the isolated-roundabout scenarios, the downstream signalized intersection was removed from the operational influence of the roundabout while maintaining the same roundabout geometry, approach distribution, and traffic composition. This allowed the effect of the nearby signalized intersection to be isolated from the effects of traffic demand and roundabout geometry (Figure 6).

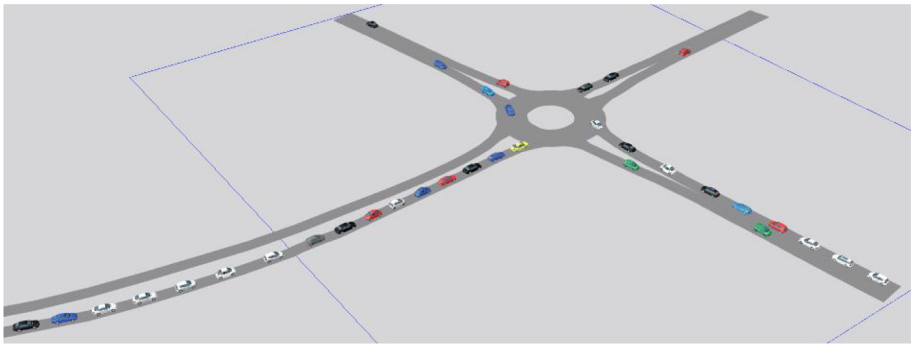


Fig. 6. VISSIM model – Scenario 4 – Isolated roundabout (ADT 20,000 veh/day)

The value of 20,000 vehicles per day was adopted as the maximum recommended Average Daily Traffic (ADT) for roundabout of this size according to the Croatian guidelines [9]. Microsimulations were conducted over a 24-hour period, with 10 simulation runs performed to ensure the reliability of the results.

The traffic performance analysis was based on several measures obtained from the VISSIM simulation model.

Average Queue Length represents the mean queue length observed during the simulation period and serves as an indicator of typical congestion levels at the analyzed location. Maximum Queue Length identifies the highest recorded queue length and is particularly relevant for detecting peak congestion and potential spillback effects.

Level of Service (LOS) represents a qualitative classification of traffic operating conditions based on vehicle delay. It ranges from LOS A, indicating free-flow or very favorable operating conditions with low delays, to LOS F, indicating oversaturated conditions with high delays and poor traffic performance [13, 14].

Vehicle Delay expresses the average time loss per vehicle relative to ideal travel conditions, that is, conditions without interference from other vehicles, traffic signals, or stopping events. As such, it directly quantifies the operational inefficiency experienced by users. Stop Delay specifically measures the portion of this delay caused by stopping.

Travel Time represents the average time required for every vehicle to traverse the analyzed route segment from the roundabout to the signalized intersection.

3 Results

The key traffic indicators obtained from the microsimulation analysis for all four scenarios are presented in Table 4. The comparison includes the roundabout–signalized intersection configuration (RB-SI) and the isolated roundabout configuration (Isolated RB), each under existing traffic demand and under an increased demand level of 20,000 veh/day. The reported VISSIM measures include Average Queue Length (QLen), Maximum Queue Length (QLenMax), Level of Service (LOS, A-F, where A represents free-flow and F represents oversaturated conditions), Vehicle Delay (VehDelay), Stop Delay, and vehicle travel time from roundabout to signalized intersection (Travel Time).

Table 4. Traffic indicators obtained from the VISSIM for the analyzed scenarios

	QLen	QLen Max	LOS	Veh Delay	Stop Delay	Stops	Travel Time
	m	*m*		*s/veh*	*s/veh*		*s*
RB-SI	4.91	88.69	C	17.14	8.05	1.17	60.33
Isolated RB	5.32	88.09	B	13.85	5.76	1.03	54.31
RB-SI_20k	72.42	266.3	F	79.30	42.42	5.16	68.12
Isolated RB_20k	71.71	257.68	F	78.28	37.90	5.62	60.59

The simulation results indicate that, under existing traffic demand, the isolated roundabout outperformed the roundabout–signalized intersection interaction scenario. In the interaction scenario (RB-SI), the average vehicle delay was 17.14 s/veh, travel time was 60.33 s, and the system operated at LOS C. In contrast, the isolated roundabout scenario produced lower delay (13.85 s/veh) and shorter travel time (54.31 s), with an improved LOS B. This corresponds to a 19.2% reduction in vehicle delay and a 10.0% reduction in travel time. Stop delay was also reduced by 28.4%, while the average number of stops decreased by 12.0%. These findings suggest that the nearby signalized intersection adversely affects roundabout performance under normal demand conditions.

When traffic volume was increased to 20,000 veh/day, both configurations operated under oversaturated conditions and reached LOS F. In the interaction scenario, vehicle delay increased to 79.30 s/veh and travel time to 68.12 s, whereas the isolated roundabout scenario yielded slightly lower values of 78.28 s/veh and 60.59 s, respectively. Relative to the interaction scenario, the isolated roundabout reduced vehicle delay by 1.3%, travel time by 11.0%, and stop delay by 10.7%. However, the average number of stops increased by 8.9%, indicating that although vehicles experienced more interruptions, the overall time spent in the system was still lower.

Comparison of the selected traffic indicators is shown in Fig. 7.

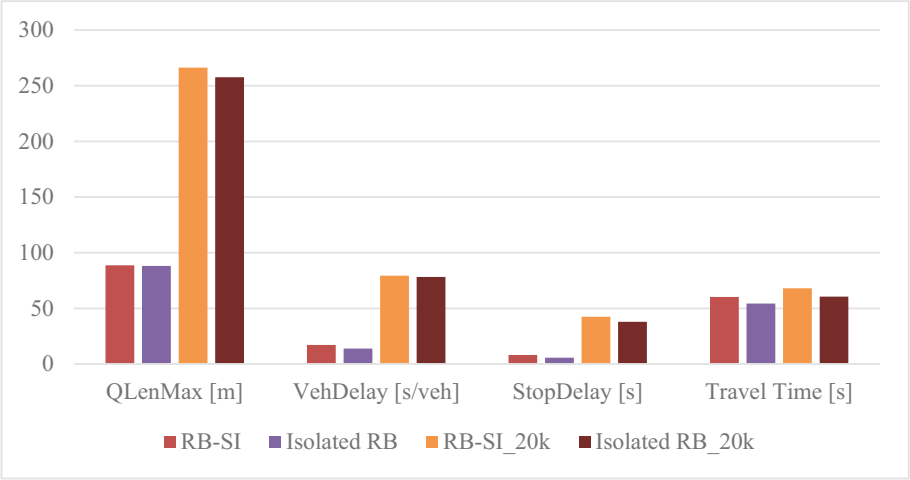


Fig. 7. Comparison of selected traffic indicators for the analyzed scenarios under current and increased traffic demand

Overall, the results show that eliminating the interaction between the roundabout and the nearby signalized intersection improves operational performance under current traffic conditions. However, when traffic demand approaches 20,000 veh/day, both layouts experience substantial deterioration, indicating that traffic demand becomes the dominant factor governing system performance.

4 Conclusion

This study examined the operational performance of a roundabout located in close proximity to a signalized intersection by using a calibrated VISSIM traffic microsimulation model based on experimentally gathered data from the City of Poreč, Croatia. Two traffic configurations were evaluated: the existing roundabout–signalized intersection interaction and a hypothetical isolated roundabout scenario. Each configuration was analyzed under both measured traffic demand and an increased demand level of 20,000 veh/day.

The results showed that, under existing traffic demand, the isolated roundabout performed more efficiently than the interacting configuration. Lower vehicle delay, lower stop delay, fewer stops, shorter travel time, and a better level of service were observed when the roundabout was not affected by the nearby signalized intersection. This suggests that closely spaced intersections can reduce the operational benefits typically associated with roundabouts, which is consistent with previous studies [8, 15, 16].

Under increased traffic demand, both configurations reached oversaturated conditions and operated at LOS F. Although the isolated roundabout still provided slightly better results in terms of delay and travel time, the differences between the two configurations became much smaller. This suggests that once traffic demand approaches the practical capacity of the roundabout, the effect of intersection interaction becomes secondary to the overall traffic volume. This is consistent with previous research [8, 17, 18], which confirms that traffic volume is the dominant factor affecting the operational

performance of a roundabout. Although the Guidelines [9] specify a maximum daily capacity of 20,000 vehicles/day for a medium-sized roundabout, in this case, with 7.1% of the daily traffic occurring during the peak hour, this traffic volume resulted in an unacceptable level of service, LOS F. These results suggest that, in addition to total daily traffic volume, the percentage of daily traffic occurring during the peak hour is also a highly influential factor, and further research is needed in this area.

The study has several limitations. It is based on a single case study with a fixed spacing of 125 m and specific geometric, traffic, and signal-control conditions, so the findings cannot be directly generalized to all urban networks. In addition, pedestrian flows and pedestrian–vehicle interactions were not analyzed, although they may affect traffic performance in urban areas.

The results highlight an important implication for urban traffic planning: the suitability of a roundabout cannot be reliably assessed using isolated-intersection analysis alone; it should be evaluated in relation to its position within the wider urban traffic network. This is especially relevant when a roundabout is located near a signalized intersection, as their interaction may significantly affect delays, queue formation, and overall corridor performance. Accordingly, corridor-level evaluation is essential, while microsimulation represents a practical tool for capturing these interdependent effects. Future research should further explore these relationships under different traffic demands, control strategies, and network settings.

Acknowledgments. This research was funded by the European Union—NextGenerationEU ("Optimizing Transport Infrastructure in the Function of Sustainable Mobility in the Cities", a scientific project at the University of Rijeka (uniri-iz-25-32) and "Safety aspects of urban transport infrastructure-SigINFRA", a scientific project at the Josip Juraj Strossmayer University of Osijek (581-UN-IOS-71)).

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

References

1. Nikiforiadis, A., Nikiforiadis, A., Mitropoulos, L., Basbas, S., Campisi, T.: International design practices for roundabouts. In: Lecture Notes in Computer Science, vol. 14111, pp. 308–326. Springer, Heidelberg (2023). https://doi.org/10.1007/978-3-031-37126-4_21/SAVE-RESEARCH
2. Forschungsgesellschaft für Straßen- und Verkehrswesen (FGSV): Richtlinien für die Anlage von Kreisverkehren. FGSV, Köln (2006)
3. Centre d'études et d'expertise sur les risques: l'environnement, la mobilité & l'aménagement (Cerema): Guide des carrefours urbains. Cerema, Lyon (2010)
4. Centrum voor Regelgeving en Onderzoek in de Wegenbouw en de Verkeerstechniek (CROW): Richtlijn rotondes (Publicatie 257). CROW, Ede (2015)
5. National Highways: Geometric design of roundabouts. National Highways, UK (2023)
6. Wisconsin Department of Transportation: Facilities Development Manual (T439). WisDOT, Wisconsin (2022)

7. Brocchini, L., Pratelli, A., Josselin, D., Losa, M.: A microsimulation-based methodology for evaluating efficiency and safety in roundabout corridors: case studies of Pisa (Italy) and Avignon (France). Infrastructures. **10**, 186 (2025). https://doi.org/10.3390/INFRASTRUCTURES10070186
8. Klobučar, M., Deluka-Tibljaš, A., Šurdonja, S., Ištoka Otković, I.: A theoretical model for optimizing signalized intersection and roundabout distance using microsimulations. Future Transp. **5**, 28 (2025). https://doi.org/10.3390/FUTURETRANSP5010028
9. Hrvatske ceste: Smjernice za projektiranje kružnih raskrižja na državnim cestama. Hrvatske ceste, Rijeka (2014)
10. Ištoka Otković, I., Tollazzi, T., Šraml, M., Varevac, D.: Calibration of the microsimulation traffic model using different neural network applications. Future Transp. **3**, 150–168 (2023). https://doi.org/10.3390/FUTURETRANSP3010010
11. Park, B., Qi, H.: Development and evaluation of a procedure for the calibration of simulation models. Transp. Res. Rec. **208**, 208–217 (2005). https://doi.org/10.1177/0361198105193400122
12. Toledo, T., Ben-Akiva, M.E., Darda, D., Jha, M., Koutsopoulos, H.N.: Calibration of microscopic traffic simulation models with aggregate data. Transp. Res. Rec. **1876**, 10–19 (2004). https://doi.org/10.3141/1876-02
13. PTV Group: PTV VISSIM 10: User Manual. PTV Group, Karlsruhe (2022)
14. Nedevska, I., Ognjenović, S., Murgul, V.: Methodology for analysing capacity and level of service for roundabouts with one lane (HCM 2000). Procedia Eng. **187**, 797–802 (2017). https://doi.org/10.1016/J.PROENG.2017.04.442
15. Hallmark, S., Fitzsimmons, E.J., Isebrands, H.N., Giese, K.: Roundabouts in signalized corridors: evaluation of traffic flow impacts. Transp. Res. Rec. **2182**, 139–147 (2010). https://doi.org/10.3141/2182-18
16. Akçelik, R.: Modeling queue spillback and nearby signal effects in a roundabout corridor. In: Proceedings of the TRB 4th International Roundabout Conference (2014)
17. Retting, R.A., Mandavilli, S., Russell, E.R., McCartt, A.T.: Roundabouts, traffic flow and public opinion. Traffic Eng. Control. **47**, 268–272 (2006)
18. Li, H., Li, J., Yang, Z., Wu, J., Chen, X.: Micro-simulation study on capacity of roundabout. In: Proceedings of the IEEE Conference on Intelligent Transportation Systems (ITSC), Washington, pp. 852–857. IEEE (2011)

Evaluating the Impact of Traffic Control System Failures on Intersection Level of Service Using Microsimulation

Denis Codroiu[1](✉), Eugen Rosca[1], Aura Rusca[1], Maja Skuric[2], Oana Dinu[1], Valentina Radu[1], and Armand Stere[1]

[1] National University of Science and Technology Politehnica Bucharest, 313 Spl Independentei, District 6, Bucharest, Romania
alexandru.codroiu@upb.ro

[2] Faculty of Maritime Studies Kotor, University of Montenegro, Put I Bokeljske brigade, 44 Kotor, Montenegro

Abstract. Traffic control systems play a critical role in ensuring an adequate level of safety at intersections for transport infrastructure users. At the same time, key performance parameters such as delay, travel time, and queue length are monitored from the design and implementation stages of these systems. However, real-world operation may lead to system malfunctions, generating negative effects on users that can be quantified through variations in the aforementioned performance parameters. The impact of such events can be assessed through analytical scenarios implemented within microsimulation models. In this study, the authors developed a simulation environment for a heavily trafficked intersection in Bucharest and generated a series of risk scenarios. The results obtained here support the development of a maintenance strategy for traffic control system components that accounts for the negative impacts induced by their failures.

Keywords: Traffic control systems · Microsimulation models · Intersection performance; Level of Service (LOS)

1 Introduction

Within urban transportation networks, signalized intersections are critical operational nodes, and traffic control systems ensure both safe efficient traffic and operation. According to the latest editions of the Highway Capacity Manual [1], the performance of these systems is typically evaluated using quantitative indicators (delays, queue length, and volume-to-capacity ratio), which form the basis for determining the Level of Service (LOS).

Recent research in the field of traffic engineering has increasingly emphasized performance monitoring and the evaluation of signalized intersections based on historical or real-time data. Traffic control systems rely on multiple hardware and software components, including detectors, controllers, communication units, and power supply systems.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 351–364, 2027.
https://doi.org/10.1007/978-3-032-30539-8_27

Failures affecting these components can alter the actual duration of the green light, disrupt the phase sequence, or cause the traffic signal to operate inefficiently.

Over the past decade, the concept of road infrastructure resilience has received increasing attention in transportation literature. Recent studies emphasize robustness, fault tolerance, and system recovery under degraded operating conditions [2]. Many of these contributions adopt a network-level or conceptual perspective, while detailed quantification at the intersection level of efficiency loss in the event of traffic control system failures remains limited.

For evaluating intersection performance under complex traffic conditions, microscopic traffic simulation has become a fundamental tool. Simulation tools, such as PTV Vissim, enable detailed representation of vehicle following behavior, lane-changing dynamics, control logic, and demand variability [3]. More recent applications of microsimulation have demonstrated its suitability for assessing the operational impact of non-standard scenarios, including demand surges, incident conditions, and traffic signal synchronization failures [4]. However, structured evaluations that integrate explicit traffic control system failure scenarios with HCM-based service level assessment remain insufficiently explored.

In this context, the present study examines the impact of traffic control system failures on intersection service levels using microsimulation techniques. A detailed simulation model was developed for a high-traffic urban intersection in Bucharest, replicating observed demand patterns and signal configurations. Several risk scenarios representing component-level failures were implemented in the simulation environment. By comparing operational indicators under normal and degraded conditions, the study quantifies the extent to which traffic control system failures affect intersection performance and supports the development of maintenance strategies based on measurable operational impacts.

Although safety is one of the fundamental objectives of traffic control systems, this study focuses on operational performance—delay, queue length, and level of service—under normal conditions and degraded operating conditions. The safety implications are beyond the scope of this paper but underscore the importance of system reliability.

The remainder of the paper is organized as follows. Section 2 reviews relevant studies on traffic control systems and microsimulation. Section 3 defines the methodology and principles for conducting the case study. Section 4 presents the simulation results for all three operational scenarios. Section 5 discusses the results, concludes the paper and presents both the conclusions and the implications.

2 Literature Review

Traffic Control Systems (TCS) constitute a strategic operational infrastructure in urban transportation networks by integrating signal controllers, detection devices, communication modules, and power supply components into a coordinated decision-making framework. Their primary objective is to ensure the safe and efficient allocation of right-of-way.

In the context of TCS, degraded performance can result from hardware failures, detection failures, communication instability, or power outages. Analyses of infrastructure interdependencies highlight that component-level failures can generate cascading

operational impacts beyond immediate hardware failures [5]. The traffic detector failures represent one of the most common causes of traffic control system degradation. Empirical research on detection accuracy and anomaly identification demonstrates that compromised detector inputs significantly influence operational efficiency [6].

Traffic flow modeling has evolved from continuous macroscopic approaches to detailed microscopic behavioral representations. The kinematic wave theory introduced by Lighthill and Whitham (1955) [7] and Richards (1956) [8] established the theoretical foundations for describing the propagation of congestion and the dynamics of shockwaves in road networks. Microscopic modeling approaches, including the vehicle tracking model proposed by Gipps (1981) [9], allow for the explicit simulation of individual vehicle dynamics under stop-and-go conditions. Comprehensive reviews of traffic flow modeling frameworks are provided by Treiber and Kesting (2013) [10], who emphasize the importance of microscopic approaches in capturing nonlinear interactions, stochastic arrival models, and the behavior of queue clearance. Although traditional approaches to traffic modeling remain relevant, recent studies show a noticeable evolution. These have focused on integrating intelligent data collection systems and digital tools into traffic analysis frameworks. Such approaches help improve data accuracy and enable a more detailed and realistic modeling of traffic behavior [11, 12].

Numerous studies validate microscopic models for evaluating traffic signal performance and control strategies, including the analysis of delays and queue lengths at signalized intersections [13, 14]. Furthermore, the integration of advanced digital technologies such as digital twins enables dynamic assessments of the performance and efficiency of traffic control infrastructure [15, 16]. Simulation-based analyses have been successfully applied in many studies to assess the impact of infrastructure changes and sustainable mobility solutions, such as Park-and-Ride systems and traffic allocation algorithms [17, 18]

Despite advances in simulation-based analysis, most studies focus on optimization, capacity improvement, or demand variability, rather than on the explicit modeling of failures in traffic control system components. The combined integration of (i) TCS reliability degradation, (ii) LOS assessment, and (iii) microscopic scenario modeling remains underdeveloped.

The literature review initially involved querying the Google Scholar database using Harzing's Publish or Perish (PoP) software, version 8. Studies written in English and accessible to the general public were selected, with no time restrictions. In order to capture the main research directions related to the topic, the search strategy employed a combination of relevant keywords, including "traffic control systems", "microsimulation", "intersection performance", "Level of Service (LOS)", and "traffic flow". The bibliometric trend illustrated in Fig. 1 confirms a substantial increase in research output related to traffic control systems [19, 20], intersection performance [21], and simulation methodologies [22–24] over the past decades.

The more extensive applicability of traffic analysis based on data can also be demonstrated through intelligent traffic monitoring applications in the logistics sector and in port access areas [25].

The accelerated growth after 2015 reflects advances in the fields of intelligent transportation systems, data analysis, and computational modeling. However, the lack of an

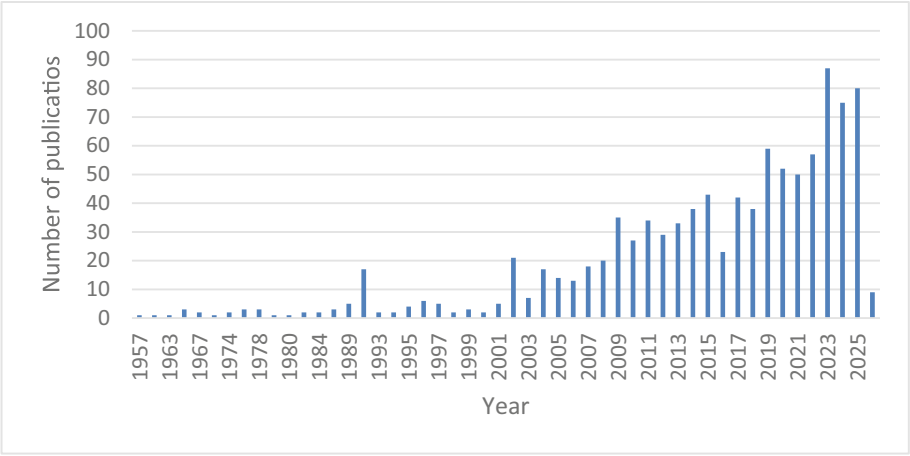


Fig. 1. Annual distribution of scientific publications related to traffic control systems, intersection performance, and microsimulation (1957–2025)

integrated framework that translates traffic control system failures into measurable LOS degradation at the component level remains insufficiently explored.

3 Methodology and Case Study

The experimental design of the study is focused on assessing the impact of the traffic control system on the operational performance of an urban intersection equipped with traffic lights, using a microsimulation-based approach. To this end, analysis scenarios were defined by modifying the parameters governing traffic control and vehicle travel behavior, under conditions of constant traffic demand.

Three operational configurations are considered: (i) standard operation of the traffic signal system, (ii) the emergency mode characterized by flashing yellow lights, which involves the elimination of traffic signal control and the occurrence of conflicting interactions between traffic flows, and (iii) traffic conditions specific to the cold season, implemented by reducing speeds and adjusting travel parameters.

The microsimulation model is configured by defining a detailed set of input parameters specific to the PTV Vissim environment, which describe both traffic demand and vehicle behavior. Traffic demand is entered in the form of input flow volumes at each access point (veh/h), distributed across travel directions (left, straight, right).

Vehicle behavior is modeled using the Wiedemann 99 algorithm by adjusting specific parameters such as CC0 (stopping distance), CC1 (headway), CC2 (headway variation), as well as dynamic acceleration and deceleration parameters. These parameters have a direct influence on the formation and dissipation of queues, evacuation capacity, and traffic flow stability.

Traffic signal timing (green, yellow, and red), cycle duration, and the way signal groups are structured are all operational parameters of the traffic control system. The saturation flow and the effective capacity of the intersection are determined by these

parameters. The values that were used helped calibrate the model against empirical data and represent actual operating conditions.

The Level of Service (LOS) is calculated using the average vehicle delay evaluating the intersection's operational performance. The Highway Capacity Manu-al (HCM) established the criteria for classification (Table 1). Delay values were obtained from the simulation, and LOS served as a benchmark metric for comparing the scenarios examined [1].

Table 1. LOS criteria for signalized intersection (source: [1])

LOS	A	B	C	D	E	F
Control Delay per vehicle (s/veh)	≤10	>10–20	>20–35	>35–55	>55–80	>80

The case study analyzes the intersection of Iuliu Maniu Blvd., Virtuții Road, and Lujerului Street, located in Sector 6 of Bucharest (Fig. 2). It is a signalized intersection classified as a complex urban intersection, characterized by the intersection of a Category I main artery with secondary distribution roads. Iuliu Maniu Boulevard is part of the European E81 network and has multiple lanes in each direction, serving as a major east-west penetration artery. Virtuții Street and Lujerului Street are Category II arteries that allow traffic to flow from north to south and connect with urban networks.

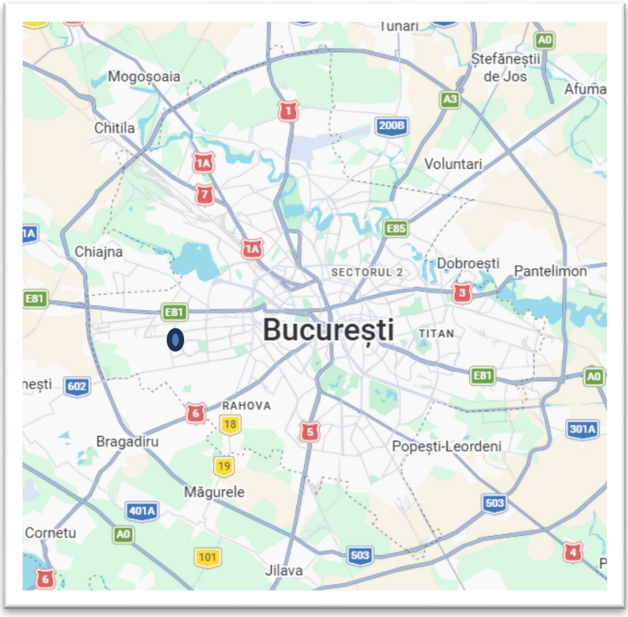


Fig. 2. Location of the analyzed intersection in Bucharest

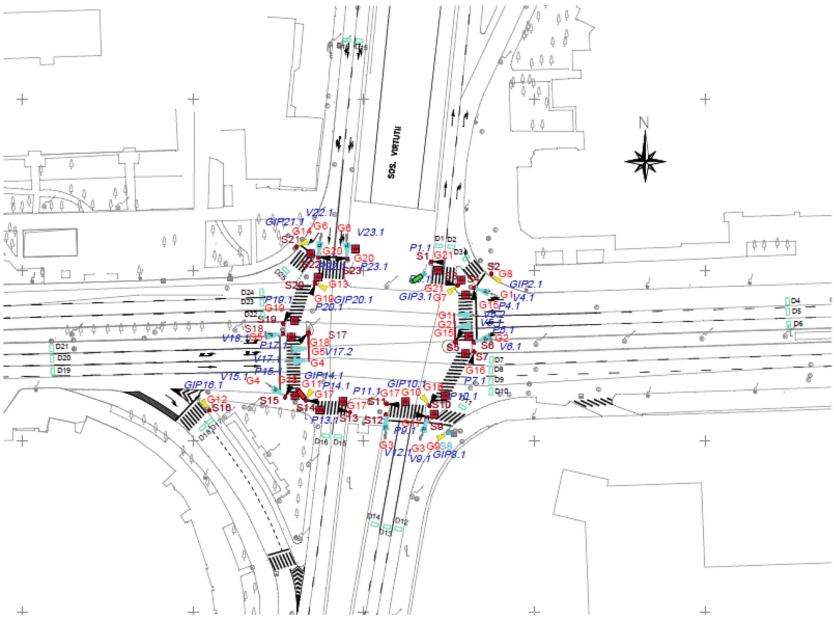


Fig. 3. Signal group configuration and induction loop detector placement (source: authors).

From a geometric standpoint, the intersection has a complex configuration, with multiple lanes on each approach and dedicated lanes for the main travel modes: left turn, straight-through, and right turn. The analysis is limited to traffic at the signalized intersection, excluding the flows associated with the Lujerului overpass, which are not considered in the simulation model. The east–west direction, corresponding to Iuliu Maniu Blvd, is the dominant direction in terms of traffic volume, leading to high demand on the intersection's capacity. To analyze traffic, directional counts were conducted at the intersection under study using inductive loop detectors placed at the entry and exit points of each access road (Fig. 3). The detection system enabled continuous recording of traffic flows for each travel route.

Traffic data were collected over a period of one week, in February, with 24-h daily coverage. For the subsequent analysis, February 9 was selected as the reference day, being considered representative of typical urban traffic conditions. From the daily recording interval, the time period 07:00–08:00 was chosen, corresponding to the morning peak hour, characterized by high traffic flow values and an increased level of intersection load. The data thus obtained formed the basis for defining the traffic demand in the microsimulation model and for its calibration in relation to real operating conditions. Traffic flows were defined for each approach of the intersection and for each movement relationship, namely left turn, through movement, and right turn, based on the data collected from the field.

The directional distribution (Table 2) highlights the predominance of movements along the east–west direction, corresponding to Iuliu Maniu Boulevard, where the highest traffic volumes are recorded. The north–south direction plays an important role in the

redistribution of flows at the local level, although the volumes are lower. Traffic control at the intersection is managed by a traffic signal system organized into six signal groups, corresponding to compatible traffic flows. The signal plan includes defined green, yellow, and red-light durations for each phase, as well as a cycle time adapted to traffic demand. The main west–east and east–west directions, corresponding to Iuliu Maniu Blvd., have longer green times due to the high traffic volumes on these routes, while the secondary directions are served by shorter green intervals.

Table 2. The simulation input data

Access	Direction	Turn	Direction ID	Percentage [%]	Volume [veh/h]	Total [veh/h]
Blvd. Iuliu Maniu (East)	Straight	→ Blvd. Iuliu Maniu (West)	D01	85%	1870	2200
	Right	→ Rd. Virtuții	D02	5%	110	
	Left	→ St. Lujerului	D03	10%	220	
St. Lujerului (South)	Right	→ Blvd. Iuliu Maniu (East)	D04	10%	176	
	Straight	→ Rd. Virtuții	D05	15%	264	1760
	Left	→ Blvd. Iuliu Maniu (West)	D06	75%	1320	
Blvd. Iuliu Maniu (West)	Right	→ St. Lujerului	D07	10%	195	
	Straight	→ Blvd. Iuliu Maniu (East)	D08	70%	1365	1950
	Left	→ Rd. Virtuții	D09	20%	390	
Rd. Virtuții (North)	Right	→ Blvd. Iuliu Maniu (West)	D10	10%	98	
	Straight	→ St. Lujerului	D11	20%	196	980
	Left	→ Blvd. Iuliu Maniu (East)	D12	70%	686	

Three operating scenarios were developed to evaluate the intersection's performance. Each scenario reflects different conditions of the traffic control system and the traffic environment. To compare the results, each scenario used the same traffic flow values. The scenarios allow for a comparison of the effects on performance indicators, particularly delay and level of service (LOS).

- Scenario 1 – Normal Operation: reflects standard intersection operation with the traffic signal system active
- Scenario 2 – Flashing Yellow: describes a degraded operating condition in which the traffic signal system is not active, and traffic flows based on right-of-way rules.
- Scenario 3 – Winter conditions: reflects a situation characterized by reduced vehicle speeds and adjusted behavior parameters, to simulate reduced traction and longer stopping distances.

Five key indicators were selected to evaluate the intersection's performance: average vehicle delay, queue length, number of stops, travel time, and LOS. These indicators provide a comprehensive analysis of the intersection's operation in terms of operational efficiency, congestion levels, and the quality of service provided to users.

4 Results

Following the simulation of the three scenarios presented above, for a cycle interval of 3600 s. To account for the stochastic nature of microscopic-level simulation models and to ensure the statistical stability of the results, a total of five simulation runs were performed for each scenario. The following values of the service parameters were obtained: maximum travel time, average travel time, average ($\overline{d}_{\text{dir}}$) and maximum delay ($d_{\max}$), as well as the average ($\overline{L}_{\text{q}}$) and maximum ($L_{\text{q}}^{\max}$) queue length, along with the number of stops (N_{stops}). From Figs. 4 and 5, it can be observed that in Scenario 2, the maximum and average travel time are improved on the priority roads within the intersection, while for the others these parameters show a clear degradation of up to 16 times in the case of maximum travel time, respectively up to 13 times in the case of average travel time. The same observations can be made for the average and maximum delay, where variations of up to 15 times can be observed (Figs. 6 and 7).

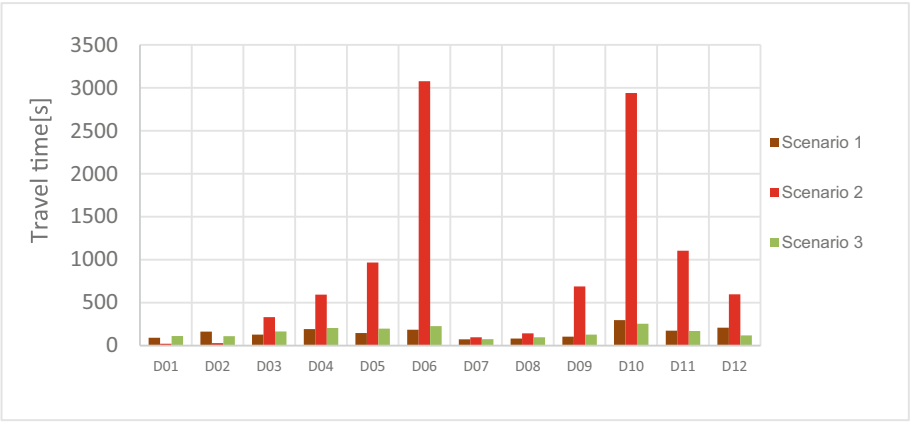


Fig. 4. Maximum Travel Time per direction and scenario

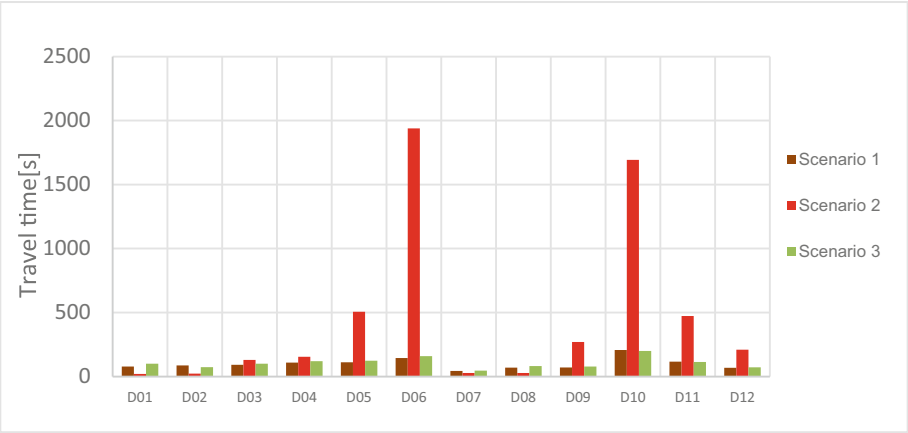


Fig. 5. Average Travel Time per direction and scenario

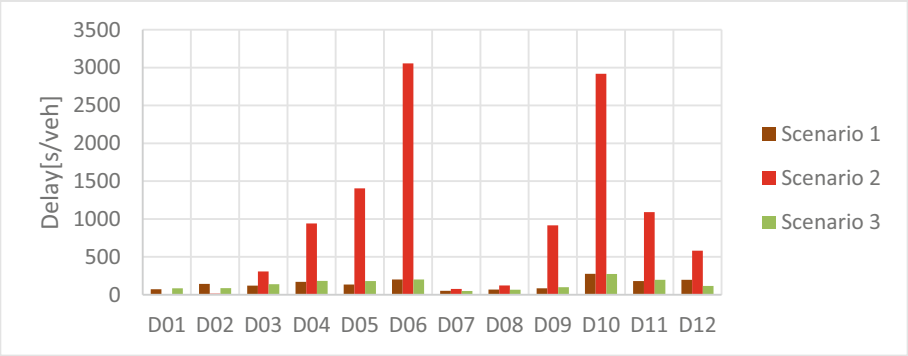


Fig. 6. Maximum delay per approach (s/veh)

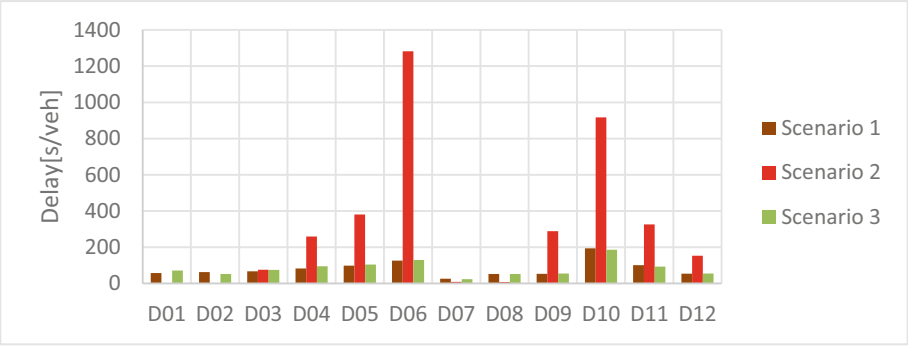


Fig. 7. Average delay per approach (s/veh)

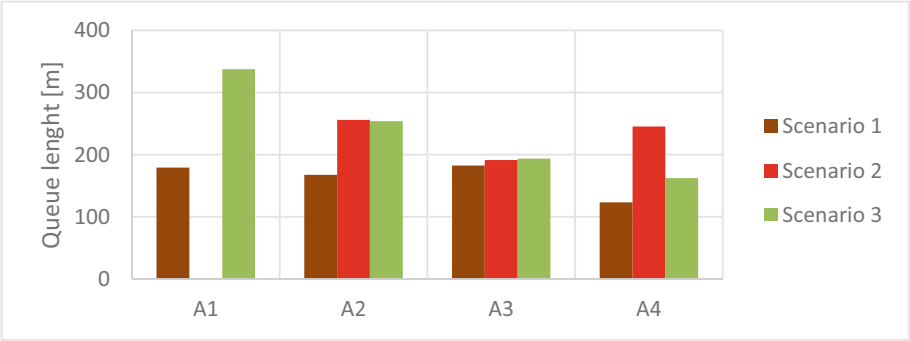


Fig. 8. Average queue length

In the case of the average queue length, as shown in Fig. 8, an increase is observed in Scenario 2, as expected, with the exception of the eastern direction, which, due to the low number of entities intending to turn left, sees most of the traffic moving along the priority east–west direction. Furthermore, the negative impact of adverse weather conditions, which leads to a reduction in vehicle speed at the intersection, is quantified by an increase in queue length in Scenario 3. The numerical values for the average and maximum queue length are presented in Table 3.

Table 3. Queue Length by Movement and Scenario

		Scenario 1			Scenario 2			Scenario 3		
Entry	Entry ID	$\bar{L}_q$ (m)	L_q^{max} (m)	N_{stops}	$\bar{L}_q$ (m)	L_q^{max} (m)	N_{stops}	$\bar{L}_q$ (m)	L_q^{max} (m)	N_{stops}
Blvd. Iuliu Maniu (East)	A1	179.22	248.18	6234	0.88	52.24	34	337.6	507.71	9224
St. Lujerului (South)	A2	167.53	184.25	4094	256.04	279.83	683	253.95	282.26	5611
Blvd. Iuliu Maniu (West)	A3	182.55	281.32	7056	191.48	406.34	4578	193.69	289.06	7359
Rd. Virtuții (North)	A4	123.17	141.57	1609	245.28	270.94	925	162.24	184.54	2240

Analyzing the intersection at a global level, regardless of the scenario, a Level of Service (LOS_{int}) E is identified, according to HCM standards (Table 1). However, as the level of analysis increases down to the level of each road converging into the intersection, it can be observed that this level (LOS_d) is not uniform across all approaches of the intersection. (Table 4).

Table 4. Intersection Level of Service by Direction and Scenario

	Scenario 1			Scenario 2			Scenario 3		
Direction ID	$\overline{d}_{dir}$ [s/veh]	LOS_d	LOS_{int}	$\overline{d}_{dir}$ [s/veh]	LOS_d	LOS_{int}	$\overline{d}_{dir}$ [s/veh]	LOS_d	LOS_{int}
D01	57.48	E	E	0.8	A	E	71.06	E	E
D02	63.06	E		2.12	A		52.51	D	
D03	66.7	E		75.49	E		74.85	E	
D04	82.46	F		259.26	F		94.69	F	
D05	98.07	F		380.63	F		104.14	F	
D06	126.04	F		1282	F		129.93	F	
D07	26.09	C		8.48	A		23.67	C	
D08	52.42	D		7.69	A		52.73	D	
D09	53.52	D		288.63	F		54.53	D	
D10	194.16	F		917.02	F		186.23	F	
D11	101.12	F		326.2	F		93.16	F	
D12	54.25	D		153.49	F		54.75	D	

Additionally, on the priority approaches within the intersection, the level of service in Scenario 2 may experience improvements, as is the case for the east–west movement direction. For Scenario 3, no variations are observed compared to the base Scenario 1.

5 Discussion and Conclusions

The simulation results show significant variability in operational parameters between traffic directions when traffic control system failures occur. In the flashing yellow light scenario (Scenario 2), the lack of traffic signal coordination leads to a significant decrease in performance on secondary access roads. This is particularly evident on directions D06 and D10, where average delays increased by over 900 s/veh compared to normal operation. These flows intersect directly with conflicting flows along the main east-west corridor, which explains the observed congestion.

In contrast, the directions along the main boulevard (D01, D03, D07, D09) show a slight improvement in traffic flow, due to the lack of restrictions regarding right-of-way. However, these advantages come at a cost, negatively affecting the efficiency of crossflow. This asymmetric behavior highlights the need for intersection-level regulation to balance the allocation of right-of-way.

It is interesting to note that, despite significant fluctuations in the directions, the overall LOS of the intersection remained constant at level E in all scenarios. This is because the HCM method combines the average control delay for all approaches, which masks the imbalances between directions. Therefore, the intersection-level LOS should be supplemented with performance indicators specific to each approach to highlight the operational effects of system degradation.

Scenario 3, which simulates winter conditions, shows a moderate decline in performance for all directions. Queue lengths increase due to reduced speeds and greater vehicle spacing, but not as much as they would increase due to the lack of traffic signal control. The operational degradation caused by environmental conditions remains more uniform and predictable than in the case of a control system failure, a fact confirmed by the simulation results.

Traffic-light-controlled intersections are a key component of urban transportation networks, as they are the points where both capacity and safety requirements converge. The analysis conducted in this study shows that changes in the traffic control system's operating mode, whether due to malfunctions or external conditions, can lead to significant variations in intersection performance.

Based on the microsimulation model developed in PTV Vissim, several operationally relevant scenarios were evaluated. The results highlight that, in the case of intermittent yellow light operation, the effects are not evenly distributed among traffic flows. The main directions may benefit from smoother traffic flow, but this advantage is offset by a marked deterioration in traffic conditions on the secondary directions, where congestion and significant delays occur. This disparity underscores the unbalanced nature of the intersection's operation in the absence of traffic signal control.

In contrast, the scenario simulating winter conditions leads to a general decline in performance, affecting all access points. Reduced speeds and changes in driver behavior result in a decrease in the intersection's evacuation capacity and, consequently, an increase in delays and queue lengths.

An important aspect highlighted by this research is the utility of microsimulation in analyzing situations that cannot be directly investigated under real-world conditions. The model allows for the reproduction of disruptive scenarios and provides a basis for quantifying their impact on operational indicators, in correlation with the HCM methodology.

From an applied perspective, the results can support the decision-making process in traffic infrastructure management by conducting ex-ante assessments of risk scenarios and identifying appropriate intervention measures. At the same time, the study highlights the need to integrate additional elements into simulation models, particularly those related to human intervention. In practice, traffic control by agents can significantly alter the behavior of traffic flows, an aspect not captured in the current form of the model.

This study also has several structural limitations. Human intervention (such as police traffic control), pedestrian flows, and adaptive traffic signal control were not modeled at this stage. The study assumed that traffic demand is constant across all scenarios, and sudden spikes in demand were not taken into account. Furthermore, this case study focuses only on operational indicators, without a direct quantification of the impact on safety.

Future work will extend the current study to include adaptive and resilient control mechanisms. Dynamic demand modeling and the inclusion of human-controlled traffic management in the event of failures will also be integrated. The use of real-time data and feedback from connected vehicles could enhance the robustness of predictive microsimulation models.

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

References

1. Transportation Research Board & National Academies of Sciences, Engineering, and Medicine: Highway Capacity Manual 7th Edition: A Guide for Multimodal Mobility Analysis. 26432. National Academies Press, Washington, D.C. (2022). https://doi.org/10.17226/26432
2. Leobons, C.M., Gouvêa Campos, V.B., Mello Bandeira, R.A.D.: Assessing urban transportation systems resilience: a proposal of indicators. Transp. Res. Procedia. **37**, 322–329 (2019)
3. Fellendorf, M., Vortisch, P.: Microscopic traffic flow simulator VISSIM. In: Barceló, J. (ed.) Fundamentals of Traffic Simulation, vol. 145, pp. 63–93. Springer, New York (2010)
4. Dakic, I., Mladenovic, M., Stevanovic, A., Zlatkovic, M.: Upgrade evaluation of traffic signal assets: high-resolution performance measurement framework. Promet Traffic Transp. **30**, 323–332 (2018)
5. Ouyang, M.: Review on modeling and simulation of interdependent critical infrastructure systems. Reliab. Eng. Syst. Saf. **121**, 43–60 (2014)
6. Yuan, T., et al.: A risk evaluation framework in system control subject to sensor degradation and failure. Sensors. **24**, 1550 (2024)
7. Lighthill, M.J., Whitham, G.B.: On kinematic waves II. A theory of traffic flow on long crowded roads. Proc. R. Soc. Lond. Ser. Math. Phys. Sci. **229**, 317–345 (1955)
8. Richards, P.I.: Shock waves on the highway. Oper. Res. **4**, 42–51 (1956)
9. Gipps, P.G.: A behavioural car-following model for computer simulation. Transp. Res. Part B Methodol. **15**, 105–111 (1981)
10. Treiber, M., Kesting, A.: Traffic Flow Dynamics: Data, Models and Simulation. Springer Berlin Heidelberg, Berlin, Heidelberg (2013). https://doi.org/10.1007/978-3-642-32460-4
11. Rusca, F., Roşca, E., Rusca, A., Ilie, A., Dinu, O.: Using smart equipment to collect data required in developing sustainable urban mobility plans. In: Computational Science and Its Applications – ICCSA 2025 Workshops Part VII, vol. 15892, pp. 101–112 (2026)
12. Opiteanu, R.-M., Rusca, F.: Development of a digital tool for quantifying and classifying traffic entities. Rom. J. Transp. Infrastruct. **14**(2), 27–27 (2025)
13. Ziemska-Osuch, M., Osuch, D.: Modeling the assessment of intersections with traffic lights and the significance level of the number of pedestrians in microsimulation models based on the PTV Vissim tool. Sustainability. **14**, 8945 (2022)
14. Hariani, M.L., et al.: Evaluation of signalized intersection performance using vissim microsimulation and PKJI 2023 (Case study: Gunung sari intersection in Cirebon City). J. Green Sci. Technol. **9** (2025)
15. Afshari, A., Lee, J., Besenski, D., Dimitrijevic, B., Spasovic, L.: Calibrating microscopic traffic simulation model using connected vehicle data and genetic algo-rithm. Appl. Sci. **15**, 1496 (2025)
16. Afshari, A., Lee, J., Besenski, D.: A digital twin platform for real-time Inter-section traffic monitoring, performance evaluation, and calibration. Infrastructures. **10**, 204 (2025)
17. Michailidis, P., Michailidis, I., Lazaridis, C.R., Kosmatopoulos, E.: Traffic signal control via reinforcement learning: a review on applications and innovations. Infrastructures. **10**, 114 (2025)

18. Rusca, F., Rusca, A., Roșca, E., Roșca, M., Dinu, O., Ghionea, F.: Algorithm for traffic allocation when developing park-and-ride facilities. In: Interdisciplinarity in Engineering (INTER-ENG 2018), vol. 32, pp. 936–943 (2019)
19. Rusca, F., Roșca, E., Rusca, A., Roșca, M., Burciu, S. Park and ride solution: a method to promote sustainable mobility. In: Proceedings of the 7th International Conference of Management and Industrial Engineering (ICMIE 2015), Bucharest, Romania, pp. 755–764 (2015)
20. Hu, Y., Rey, D., Mohajerpoor, R., Saberi, M.: Optimizing traffic signal control for continuous-flow intersections: benchmarking against a state-of-practice model. IET Intell. Transp. Syst. **18**, 2152–2165 (2024)
21. Zare, N., Tumminello, M.L., Macioszek, E., Granà, A.: Beyond efficiency: integrating resilience into the assessment of road intersection performance. Smart Cities. **8**, 184 (2025)
22. Syla, V., Lala, A.: Microsimulation-based evaluation of fixed-time and vehicle-actuated traffic signal control strategies using PTV Vissim in an urban corridor. Informatica. **50** (2026)
23. Maciejewski, M.: A comparison of microscopic traffic flow simulation systems for an urban area. Transp. Prob. Int. Sci. J. **5** (2010)
24. Roșca, M., Roșca, E., Oprea, C., Rusca, F., Dinu, O., Rusca, A.: The effects of new infrastructure on traffic dynamics: an urban simulation-based model. In: Interdisciplinarity in Engineering (INTER-ENG 2024), vol. 386, pp. 235–246 (2024)
25. 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: Advances in Maritime Technology and Engineering (MARTECH 2024), vol. 1, pp. 81–86 (2024)

Control Point Detection for Network Alignment Using Large Language Model

Müslüm Hacar(✉) and Özge Öztürk Hacar

Yildiz Technical University, Esenler, Istanbul 34210, Türkiye
mhacar@yildiz.edu.tr

Abstract. Control point (CP) selection is a crucial preprocessing step in network alignment workflows in GIS and spatial data integration. CP-based registration commonly depends on selecting corresponding intersections between a source network and a target network, after which a geometric transformation such as similarity, affine, projective, or locally adaptive warping can be estimated. This paper evaluates GPT-5.2 "thinking" mode, a multimodal Large Language Model (LLM) for CP detection and correspondence establishment between two network inputs that represent the same geographic extent. We analyze twelve experimental runs organized by three factors: input modality as rasterized network imagery versus vector networks, availability of explicit intersection node markers and identifier labels versus line-only inputs, and distortion level of the target network including regular affine-like distortions, globally irregular distortions, and local distortions. The results indicate that regular affine-like deformations enable strong LLM-assisted CP detection, particularly when nodes are provided and the raster/vector representation is chosen appropriately. This advantage diminishes under irregular distortions (global or local), where achieving consistently good performance becomes more difficult in vector-based settings. Finally, when deformation exists by localized detail-level differences, neither raster nor vector yields reliable results, making local distortions a modality-agnostic limitation for the current LLM-assisted CP detection setup.

Keywords: Large Language Model · Alignment · Georeferencing · Road Network · Coordinate Transformation · GeoAI

1 Introduction

Map alignment is central to a wide range of geospatial tasks, including map conflation of multi-source data, integration of utility graphs, map updating, and harmonization of digitized, aerial-based, remotely sensed, and surveyed datasets. When two maps represent the same geographic region but differ by georeferencing error or deformation, a common practical approach is to select corresponding control points (CPs) and estimate a transformation that aligns one map to the other. Corners of map sheets, intersections at cartesian grid networks, and junctions at street networks are natural CPs since they tend to be semantically stable and often distinctive in local structure. The identification and matching of such points remains manual, especially when representations differ across raster and vector formats or when distortions are spatially complex.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 365–373, 2027.
https://doi.org/10.1007/978-3-032-30539-8_28

CP-based registration and georeferencing has long been a standard mechanism in GIS and cartography, relying on human selection of tie points and the fitting of parametric transforms such as similarity and affine transformations or higher-order and spline-based warps when distortions are complex. In network contexts, intersections are favored CPs since they represent topological events where multiple edges meet, and their local neighborhood can, in principle, provide additional cues beyond raw coordinates. Nevertheless, CP selection remains a manual step in many production environments, particularly in cases of representation mismatch or local deformation where geometric heuristics break down.

Automated network conflation and graph matching methods address the same underlying problem by attempting to compute correspondences between edges, nodes, or subgraphs [1–5]. These approaches range from geometry-driven techniques that use proximity, overlap, and distance measures to topology-aware methods that incorporate degree distributions, adjacency patterns, and subgraph signatures. While these methods can be effective, they often require careful tuning and may be sensitive to missing edges, inconsistent sampling density, or distortions that vary spatially. Local distortion can defeat nearest-neighbor matching and can complicate the use of a single global model for correspondence.

In parallel, computer vision and point-set registration offer families of methods for matching under noise and deformation, including robust estimation and outlier rejection [6–10]. Their performance, however, depends on the availability of stable features in both input maps (or spatial data) and on whether the deformation difference can be captured by the chosen model. When raster linework is warped or aliased, key point salience can degrade and the definition of a "point" representation for junctions becomes ambiguous.

Multimodal LLMs are increasingly used for visual reasoning tasks that involve associating symbols, reading labels, and describing spatial relationships in images and diagrams. Their potential contribution to GIS workflows was evaluated recently in the studies [11–21] lies less in replacing classical geometric computation and more in providing interactive assistance, triaging candidate matches, and adapting to heterogeneous representations through language-guided reasoning. However, the reliability of such models in CP detection and correspondence under geometric distortion has not been widely established. This paper contributes empirical evidence for that setting, emphasizing the conditions under which LLM assistance should be constrained and combined with classical candidate generation. In this study, we use a multimodal LLM, GPT-5.2 "thinking" mode in a zero-shot setting, developed by OpenAI [22]. This work extends our earlier assessment of LLM-based map conflation, in which GPT-5 "thinking" mode was evaluated for line-to-line road correspondence under strict 1:1 and flexible M:N matching strategies [23]. While that study addressed edge-level conflation on topologically structured networks, the present paper focuses on a junction-level CP detection and correspondence under controlled geometric distortions. This paper investigates the reliability of LLM-assisted CP detection under controlled conditions that resemble real-world data heterogeneity: different input modalities, the presence or absence of explicit node layers with identifiers, and different distortion regimes that stress both detection and matching. The study is based on twelve experimental tests runs in which the same general task is executed under different combinations of modality, node availability,

and distortion level, and the outcomes are recorded as accuracy along with qualitative observations of error behavior.

2 Study Data, Experimental Design, and LLM-Assisted CP Detection

Figure 1 defines the experimental design through six source–target input combinations used for CP detection. In every combination, the left panel represents the source (original) network and the right panel represents the target network obtained by applying a controlled deformation operator. These six combinations are treated as distinct processes because they change the underlying difficulty of the CP task by independently varying the deformation regime applied to the target and whether intersections are provided as explicit node candidates. Each of the six processes is then instantiated in two data modalities: raster and vector, yielding twelve total experimental processes.

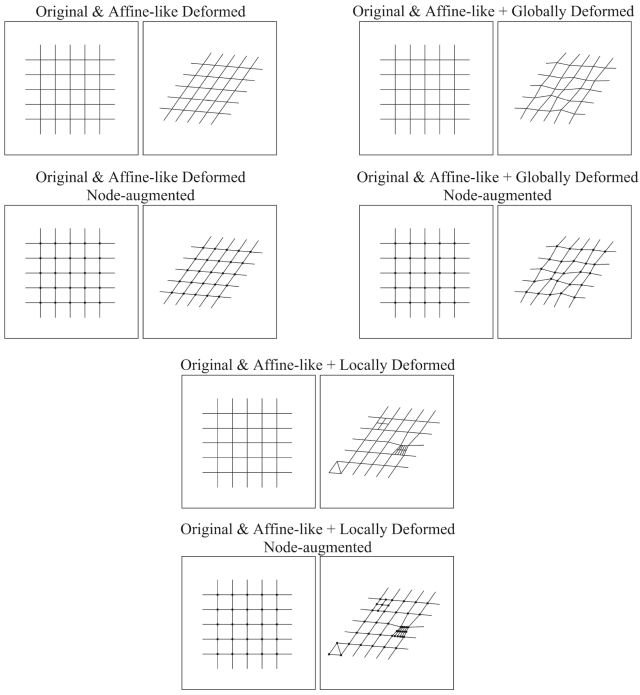


Fig. 1. Input combinations for CP detection.

The first axis of variation is the deformation regime. The "Original & Affine-like Deformed" combination represents globally consistent misalignment that can be approximated by a single parametric mapping; in practice this regime is associated with shifting, shear, scale, and rotation effects that preserve a coherent transformation across the entire

extent (Fig. 2). The "Original & Affine-like + Globally Deformed" combination represents a more challenging target in which a global affine-like component is present but additional globally acting deformation is superimposed, reducing global consistency and weakening the reliability of purely distance-based matching logic (Fig. 3). The "Original & Affine-like + Locally Deformed" combination represents spatially heterogeneous deformation concentrated within subregions; this regime is characterized by locally inconsistent displacements that may alter the apparent neighborhood structure of intersections within the deformed area while leaving other regions relatively coherent (Fig. 3). These three regimes allow the study to separate performance under globally explainable misalignment from performance under irregular and spatially varying deformation.

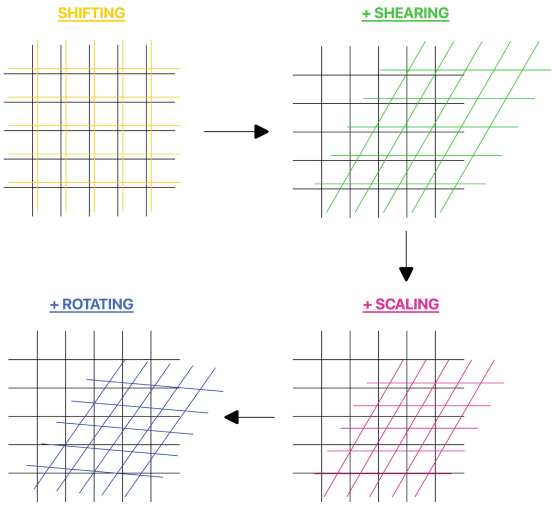


Fig. 2. Original (black), shifted (yellow), sheared (green), scaled (pink), and rotated (blue) networks.

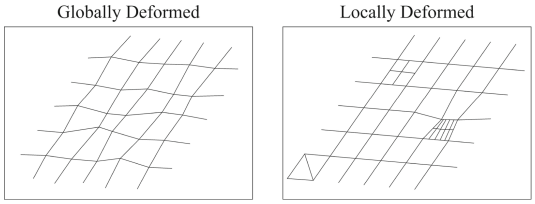


Fig. 3. Irregular deformations: Globally (left) and locally (right) distorted networks.

The second axis of variation is the availability of explicit intersection nodes. For each deformation regime, the experiment includes a line-only and a node-augmented variants (Fig. 1 and Table 1). In the line-only variant, intersections are just crossing lines and must be inferred from network geometry, which makes the task a joint detection

and matching problem. This means that the model must first decide where valid CPs might exist and then establish correspondences across the source and target. In the node-augmented variant, intersections are explicitly marked as points, which constrains the task to selecting and pairing from an explicit candidate set of junction points. Availability of nodes is assumed reducing the ambiguity in CP definition by shifting the task toward correspondence establishment under deformation. Using these variants is central to the evaluation because it isolates errors due to control point discovery from errors due to correspondence reasoning.

Table 1. Assisting prompts used to conduct experiments in LLM.

Input combination	Network structure	Assisting prompt
Original & Affine-like Deformed	Node-augmented	Possibility of one-to-one alignment At least 4 CPs
	Line-only	At least 4 CPs
Original & Affine-like + Globally Deformed	Node-augmented	Possibility of one-to-one alignment At least 4 CPs
	Line-only	N/A
Original & Affine-like + Locally Deformed	Node-augmented	At least 4 CPs
	Line-only	N/A

Each of the six processes, shown in Fig. 1, is executed in two modalities. In the raster setting, inputs are provided as network renderings in .tiff format, in which linework and, when applicable, node marks are interpreted as image structures. In this modality, intersection salience depends on pixel-level cues such as stroke crossings and junction shapes. In the vector setting, inputs are provided as line-based network geometry in .gpkg format. In node-augmented variants, an explicit point representation of intersections is provided as a separate .gpkg file. Comparisons between the processes carried out with raster and vector inputs isolate modality effects, including whether intersection salience in raster is preserved under deformation and whether vector line-only inputs introduce ambiguity in the absence of explicit node candidates.

In all experiments, the LLM was first given a set of standard task instructions requiring it to read the provided files, identify corresponding CPs for network alignment, and store the resulting CP associations in a downloadable output file. In addition to these baseline instructions, the assisting prompts listed in Table 1 were incorporated in the relevant tests to further guide the model's behavior. These auxiliary prompts were used selectively depending on the input combination and network structure, for example by explicitly framing the task as a one-to-one alignment problem in node-augmented settings and by enforcing a minimum number of CPs when appropriate.

3 Experimental Results

The results of raster tests in Table 2 and vector tests in Table 3 indicate that affine-like distortion is the regime in which LLM assistance is most reliable, particularly when explicit node markers with identifiers are available. In the node-augmented condition under affine-like distortion, both raster and vector runs achieve perfect correctness, consistent with the expectation that labeled node anchors make correspondence largely deterministic. In the line-only condition under the same distortion, raster achieves perfect accuracy for the requested minimum set of four CPs, whereas vector drops to half accuracy. This suggests that when the task is restricted to producing a small set of intersections under a globally consistent warp, raster imagery may offer cues that are easier for the LLM model to exploit than vector linework.

Table 2. Performance statistics in the CP detection scenarios using raster sources.

Input combination	Network structure	Total CP	Correct CP	%
Original & Affine-like Deformed	Node-augmented	25	25	100
	Line-only	4	4	100
Original & Affine-like + Globally Deformed	Node-augmented	25	25	100
	Line-only	20	2	10
Original & Affine-like + Locally Deformed	Node-augmented	18	9	50
	Line-only	25	8	32

Table 3. Performance statistics in the CP detection scenarios using vector sources.

Input combination	Network structure	Total CP	Correct CP	%
Original & Affine-like Deformed	Node-augmented	25	25	100
	Line-only	4	2	50
Original & Affine-like + Globally Deformed	Node-augmented	25	11	44
	Line-only	16	5	31.3
Original & Affine-like + Locally Deformed	Node-augmented	16	5	31.3
	Line-only	32	12	37.5

Under global and irregular distortion, performance diverges sharply across conditions. The experiment with node-augmented network in raster format remains perfect, while vector test drops substantially, indicating that even with labeled vector nodes the model can mis-associate correspondences when irregular deformation disrupts the spatial relationships the model expects (Tables 2 and 3). In the test with line-only rasterized network, the performance collapses to a very low accuracy and the qualitative

record indicates that the model selected control points along network edges rather than at intersections, which constitutes a primary detection failure.

Under local distortion, correctness percentages are consistently degraded across all configurations, with no single representation dominating. The qualitative observations highlight that locally distorted regions attract incorrect selections and that proximity-based matching remains a frequent default strategy, which is brittle when distortions vary spatially. Consequently, the study shows three significant results: (1) Regular affine-like deformations can yield strong LLM-assisted CP detection performance, especially when supported by explicit node usage and, where beneficial, an appropriate raster/vector representation choice. (2) the same advantage does not carry over when the discrepancy between the two inputs becomes irregular (whether globally irregular or locally irregular), where achieving consistently good performance becomes more difficult in vector-based settings. (3) When the mismatch is dominated by localized detail-level deformation differences, neither raster nor vector inputs produced reliable results. As a results, local distortions constitute a modality-agnostic limitation for the current LLM-assisted CP detection setup.

4 Discussion and Conclusion

The experiments demonstrate that LLM-assisted CP detection is strongly conditioned by both deformation regularity and the availability of explicit intersection nodes. Under regular affine-like distortions, performance is consistently high, and the node-augmented setting yields perfect correctness in both raster and vector inputs, confirming that labeled junction anchors effectively constrain the correspondence problem. When nodes are not provided, performance becomes more representation-dependent: raster linework supports accurate selection of a small minimum set of intersections under globally consistent warps, whereas vector line-only inputs show reduced correctness, suggesting that the absence of an explicit node representation increases ambiguity in how intersections are operationalized and paired.

When the discrepancy between inputs becomes irregular, the reliability observed in affine-like cases does not generalize. In particular, the correctness of vector node-augmented case decreased significantly, remarking that even with labeled nodes the LLM-assisted detection may fall back on proximity-driven reasoning and mis-associate correspondences when irregular deformation violates expected spatial relationships. In the raster line-only irregular case, correctness collapses primarily due to a detection failure, where the model selects points along edges rather than true intersections. These outcomes indicate that irregular deformation stresses both stages of the problem: identifying valid junction CPs and establishing one-to-one correspondence. In that case, the dominant error mode shifts from candidate ambiguity to systematic mis-selection and mis-association.

Local distortions produce consistently degraded correctness across all configurations, with only modest differences between raster and vector and between node-augmented and line-only settings. The qualitative pattern suggests that locally deformed regions attract incorrect selections and that proximity-based matching becomes brittle when deformation varies spatially, limiting the benefit of representation choice.

Consequently, while regular affine-like misalignments can be handled effectively through node augmentation and, in some cases, careful modality selection, irregular and especially locally varying distortions represent the principal limitation of the current LLM-assisted CP detection setup, as neither raster nor vector inputs provide consistently reliable performance. In the future, the LLM-assisted CP detection may benefit from using additional assisting prompts leading the LLM actions with topology-aware matching cues retrieved from source network structures.

Acknowledgments. This study was supported by the Scientific and Technological Research Council of Türkiye (TÜBİTAK) under the 2224-A - Grant Program for Participation in Scientific Meetings Abroad. The authors gratefully acknowledge this support.

Disclosure of Interests The author declares no conflict of interest.

References

1. Saalfeld, A.: Conflation automated map compilation. Int. J. Geogr. Inf. Syst. **2**, 217–228 (1988). https://doi.org/10.1080/02693798808927897
2. Rosen, B., Saalfeld, A.: Match criteria for automatic alignment. In: Proceedings of the 7th International Symposium on Computer-Assisted Cartography (Auto-Carto 7), Washington, D.C. (1985)
3. López-Vázquez, C., Manso Callejo, M.A.: Point- and curve-based geometric conflation. Int. J. Geogr. Inf. Sci. **27**, 192–207 (2013). https://doi.org/10.1080/13658816.2012.677537
4. Doytsher, Y., Filin, S., Ezra, E.: Transformation of datasets in a linear-based map conflation framework. Surv. Land Inf. Syst. **61**, 159–169 (2001)
5. Li, L., Goodchild, M.F.: An optimisation model for linear feature matching in geographical data conflation. Int. J. Image Data Fusion. **2**, 309–328 (2011). https://doi.org/10.1080/19479832.2011.577458
6. Besl, P.J., McKay, N.D.: A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. **14**, 239–256 (1992). https://doi.org/10.1109/34.121791
7. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM. **24**, 381–395 (1981). https://doi.org/10.1145/358669.358692
8. Bookstein, F.L.: Principal warps: thin-plate splines and the decomposition of deformations. IEEE Trans. Pattern Anal. Mach. Intell. **11**, 567–585 (1989). https://doi.org/10.1109/34.24792
9. Myronenko, A., Song, X.: Point set registration: coherent point drift. IEEE Trans. Pattern Anal. Mach. Intell. **32**, 2262–2275 (2010). https://doi.org/10.1109/TPAMI.2010.46
10. Ma, J., Zhao, J., Tian, J., Yuille, A.L., Tu, Z.: Robust point matching via vector field consensus. IEEE Trans. Image Process. **23**, 1706–1721 (2014). https://doi.org/10.1109/TIP.2014.2307478
11. Tao, R., Xu, J.: Mapping with ChatGPT. ISPRS Int. J. Geo Inf. **12**, 284 (2023). https://doi.org/10.3390/ijgi12070284
12. Mooney, P., Cui, W., Guan, B., Juhász, L.: Towards understanding the geospatial skills of ChatGPT: taking a geographic information systems (GIS) exam. In: Proceedings of the 6th ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery, New York, NY, USA, pp. 85–94 (2023)
13. Mansourian, A., Oucheikh, R.: ChatGeoAI: enabling geospatial analysis for public through natural language, with large language models. ISPRS Int. J. Geo Inf. **13**, 348 (2024). https://doi.org/10.3390/ijgi13100348

14. Zhang, Y., Wei, C., He, Z., Yu, W.: GeoGPT: an assistant for understanding and processing geospatial tasks. Int. J. Appl. Earth Obs. Geoinf. **131**, 103976 (2024). https://doi.org/10.1016/j.jag.2024.103976
15. Zhang, Y. et al.: Geospatial large language model trained with a simulated environment for generating tool-use chains autonomously. Int. J. Appl. Earth Obs. Geoinf. **136**, 104312 (2024). https://doi.org/10.1016/j.jag.2024.104312
16. Gulnerman, A.G.: AI-driven urban perception mapping: comparing Gemini and LLaVA. In: International Conference on Computational Science and its Applications, Istanbul, Türkiye, pp. 339–349 (2025). https://doi.org/10.1007/978-3-031-97663-6_30
17. Memduhoğlu, A.: Towards AI-assisted mapmaking: assessing the capabilities of GPT-4o in cartographic design. ISPRS Int. J. Geo Inf. **14**, 35 (2025). https://doi.org/10.3390/ijgi14010035
18. Pierdicca, R., Muralikrishna, N., Tonetto, F., Ghianda, A.: On the use of LLMs for GIS-based spatial analysis. ISPRS Int. J. Geo Inf. **14**, 401 (2025). https://doi.org/10.3390/ijgi14100401
19. Chen, G. et al.: GeoJSEval: an automated evaluation framework for large language models on JavaScript-based geospatial computation and visualization code generation. ISPRS Int. J. Geo Inf. **14**, 382 (2025). https://doi.org/10.3390/ijgi14100382
20. Wu, H. et al.: AutoGEEval: a multimodal and automated evaluation framework for geospatial code generation on GEE with large language models. ISPRS Int. J. Geo Inf. **14**, 256 (2025). https://doi.org/10.3390/ijgi14070256
21. Wu, H. et al.: AutoGEEval++: a multi-level and multi-geospatial-modality automated evaluation framework for large language models in geospatial code generation on Google Earth Engine. Big Earth Data. **9**, 747–796 (2025). https://doi.org/10.1080/20964471.2025.2581425
22. OpenAI: GPT-5.2 System Card Addendum (2025). https://openai.com/index/gpt-5-system-card-update-gpt-5-2/
23. Hacar, M., Öztürk Hacar, Ö.: LLM-based map conflation: performance assessment on matching embedded road lines. ISPRS Int. J. Geo Inf. **15**, 144 (2026). https://doi.org/10.3390/ijgi15040144



Active Reflectors and Passive Corner Reflectors for Deformation Monitoring Using DInSAR Interferometry: Some Remarks from Sentinel-1 Imagery Tests

G. Luzi(✉), P. Espín-López, and Q. Gao

Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA), B6/B4, Parc Mediterrani de la Tecnologia (PMT), Av. Carl Friedrich Gauss 7, 08860 Castelldefels, Spain
gluzi@cttc.es

Abstract. Passive corner reflectors are often used in spaceborne Synthetic Aperture Radar interferometry as standard ground points to achieve the monitoring of areas where Permanent Scatters are not available, as in uncoherent areas when the radar backscattering of natural targets is unstable. The main goal of the use of passive reflectors in radar interferometry is to provide a reference to check the stability of the area, for validation purposes, and often to reduce the effect of phase artefacts in data processing. The use of passive reflectors, whose size depends on the operating wavelength of the sensor, can be not suitable, especially in areas like glaciers, snow covered regions, and mountain slopes, where accessibility and installation can be very time and cost consuming, and severe weather conditions can jeopardize their performance. An alternative to passive reflectors can be the installation of active reflectors, compact and lighter, but it is mandatory to have in hands an electric power source. The stability of its amplitude and phase response is basically affected by the presence of active radio frequency components susceptible to temperature changes. When active reflectors are installed for months-to-year-long epochs, the natural variation of the scenario air temperature demands a careful understanding of the device's thermal behavior. Some products have been made available from the market for field campaigns but till now their use has not yet spread. In this paper the authors introduce the basics of passive and active reflectors and compare the recent performance obtained using an active device used for three years to estimate deformation. Finally results available from recent literature are compared to obtain the reliability and performance of the use of an active reflectors in deformation monitoring, estimating the achieved accuracy and providing suggestions for their use.

Keywords: SAR · Interferometry · Sentinel-1 · Corner Reflectors · Deformation

1 Introduction

In Radar spaceborne remote sensing Passive Corner Reflectors (PCR) are often used in Synthetic Aperture Radar (SAR) imaging, and in Differential Interferometric SAR (DInSAR) as standard ground points to increase the extension of the study areas where

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 374–383, 2027.
https://doi.org/10.1007/978-3-032-30539-8_29

Permanent Scatters (PS) are scarce or not available [1–3]. These limitations are usually due to a weak or instable temporal coherence over the survey interval or more generally to an instable radar scenario. Furthermore, the use of such reference targets can support the achievement of higher reliability in final products. The main goal of the use of PCRs in radar interferometry for deformation monitoring is to provide reference data to check the stability of the area, for validation purposes, also reducing the effect of phase artefacts due to atmospheric humidity and temperature variations. These targets appear in amplitude image as pixels with high amplitude and stable response over long temporal periods thanks to their high reflectivity, and they can also provide stable differential phase response too. Sometimes the use of PCRs, whose size depends on the operating wavelength of the sensor, is not suitable for low frequency (i.e. large wavelengths) cases as, S and P bands, and sometimes also for C band, especially in areas like glaciers, snow covered regions, and mountain slopes, where accessibility and PCRs' installation can be very time and cost consuming. Furthermore, severe weather conditions such as strong winds or snow and rain can jeopardize their performance. An alternative to PCRs can be represented by the installation of Active Reflectors (AR), compact and lighter but for which it is mandatory to have in hands a power source.

While a passive reflector simply reflects toward the satellite the satellite signal impinging on its surface using multiple reflections (see Fig. 1a), an AR, when the satellite passes overhead, receives the signal through a first antenna and retransmit it with a higher strength, using the energy of a power source to amplify its signal and making its response stronger than that received one (see Fig. 2a). The stability of its amplitude and phase response, which is fundamental for a successful employ, is basically affected by the presence of active Radio Frequency (RF) components susceptible to temperature changes and the RF circuit usually adds a delay due to electrical connections which must carefully deal with to avoid error in target localization. When ARs are installed for months to yearlong epochs, the electronic and radiofrequency circuits of these devices can be affected by the seasonal variations of the air temperature, and this issue demands a careful thermal understanding of the device's behavior. Some products have been made available from the market for field campaigns but till now their use has not yet spread for two main reasons: the challenge issue of achieving a phase stability as that of PCR and, nevertheless the lower cost of installation, their price is not drastically competitive with respect to that of passive device as discussed in [4]. In this paper we introduce the basics of PCR and AR and compare the recent results obtained using a prototype in the framework of a activity carried out to use DInSAR to estimate deformation of natural surfaces, to those from some studies available from literature. This paper is organized into four sections. After this introduction, the manuscript includes a synthetic section recalling the basic functioning principle of a passive target and of the AR rationale. Some examples from experimental campaigns carried out by the authors follow in the specific section. Main conclusions and remarks are finally discussed in the last section of the paper.

2 Reference Targets for DInSAR Processing

In SAR imaging a processed SLC (Single Look Complex) image associates an amplitude and phase value to each pixel. Differential phase information can be retrieved using two or more images, while amplitude values represent the reflectivity associated to the pixel, or more precisely the amplitude of the backscattering coefficient; its value can be associated to surfaces and to outstanding targets included in the pixel area. The amount of energy reflected by the target can be estimated through its Radar Cross Section (RCS), a parameter defined on the bases of the radar equation [5] and which quantitatively represents the area intercepting that amount of power which, if radiated isotropically, produces the same received power in the radar. Targets with an RCS are significantly higher than the signals coming from the surrounding pixels, namely the radar clutter, appear brighter in the image, easy detectable, and can provide a high phase stability. As DInSAR processing needs phase stable pixels to retrieve the surface deformation, these targets are often installed in the observed areas to monitor ground motion by techniques such as the Persistent Scattering (PS) to retrieve temporal series [6], achieving deformation estimates from a set of SAR images acquired at different dates. Deploying artificial coherent scatters in areas with low interferometric coherence as vegetation layers, which can mask the surface radar response, or snow-covered slope, or glaciers, improve the density of coherent points making the in DInSAR technique application more affordable.

2.1 The PCR

A passive corner reflector (PCR) is a device capable of providing high reflectivity when illuminated by a radar signal. These objects are typically metal structures, with a specific geometry that allows for the efficiency of a large surface area with an optimized shape for its footprint; its capacity to backscatter the impinging radar signal is due to the specular reflectivity of the metal surface at microwaves and on its internal multiple reflections as depicted in Fig. 1a. The PCR 's response is typically not very directional, as this would otherwise make it difficult to orient when installed for satellite imagery acquisition. It allows to estimate the differential phase response which is the fundamental parameter from which using DInSAR measurements can retrieve estimates of the deformation over the monitored areas. To evaluate the quantitative achievement of a target response, RCS must be sufficiently high, and the Signal to Clutter Ratio (SCR) which estimate the ratio between the RCS of the target and the signal, as noise, coming from the pixel surrounding that corresponding to the target [7]. A detailed knowledge of the response of PCR must consider factors as the geometry of the installation and the surrounding surface, as deeply analyzed in [8], but most of their applications use a basic approach to obtain its signature. In particular, the phase accuracy expected in DInSAR data processing is usually estimated through the magnitude of the SCR and for a stochastic backscattering environment, a simple formula can be used [9]:

$$\Delta\varphi = \frac{1}{\sqrt{2\,SCR}}$$

Triangle Trihedral Corner Reflectors (TTCR) are the most diffused PCR used to calibrate the amplitude response available from a SAR image and give a reference phase point in DInSAR data processing. Their main performance is a fairly large RCS for its size, over a broad range of aspect angles, whose maximum can be easy calculated through the following formula:

$$RCS = \frac{4\,\pi\,l^4}{3\,\lambda^2}$$

where λ is the radar wavelength, and l is the inner-leg dimension, RCS is in m^2.

When designed to be used for C band, the operating frequency of Sentinel-1 (S1), and even more with spaceborne radar operating at lower frequencies, the PCRs are bulky because the maximum achievable RCS is related to its size. For example, to obtain an expected accuracy of 1 mm at C band (e.g Sentinel-1 signal with centre wavelength = 5.5467 cm) is necessary to install a TCR whose size is greater than 1.7 m (see Table 1 pag.6 [9]). This fact often results in further issues when hard weather conditions, such as heavy rainfall, strong winds and snowfall are frequently occurring in the installation site. In addition, their placement can be difficult in sites of hard accessibility as mountain slopes, glaciers etc. Different kind of PCR have been proposed in literature and recently the test a novel low-cost prototype has been published [4]. In Fig. 1b a picture of a Square Trihedral Corner Reflector (STTR), a metal structure 0.60 m size, installed in a mountain area to support the monitoring of a landslide in the framework of the deformation monitoring campaign carried out using Sentinel-1 imagery and DInSAR processing [10]. This type of PCR with respect to the most used TTCR performs a higher RCS without increasing excessively the encumbrance and the installation at the cost of slightly smaller angle of aperture.

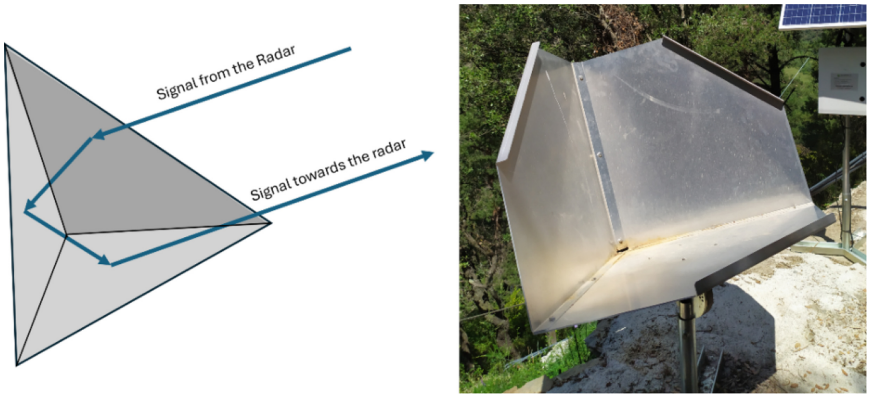


Fig. 1. Left: Working principle of a TTCR; right: Photo of a STCR.

2.2 The AR

Differently from a passive reflector, an AR, receives, amplifies and retransmit the received signal finally making its response stronger than that received one. The stability

of its amplitude and phase response, which is fundamental for a successful employ, is basically affected by the presence of active Radio Frequency (RF) components susceptible to temperature changes. In this paper we analyse the performance of an AR designed and implemented according to a design rationale focused on the achievement of the performance required for its use in deformation monitoring where aspects as a low cost, easy installation, and a low power consumption. These characteristics make simpler the setup of a dense network often demanded in large low co-herence areas deformation monitoring. In Fig. 2 the working principle and a photo of an AR prototype are shown.

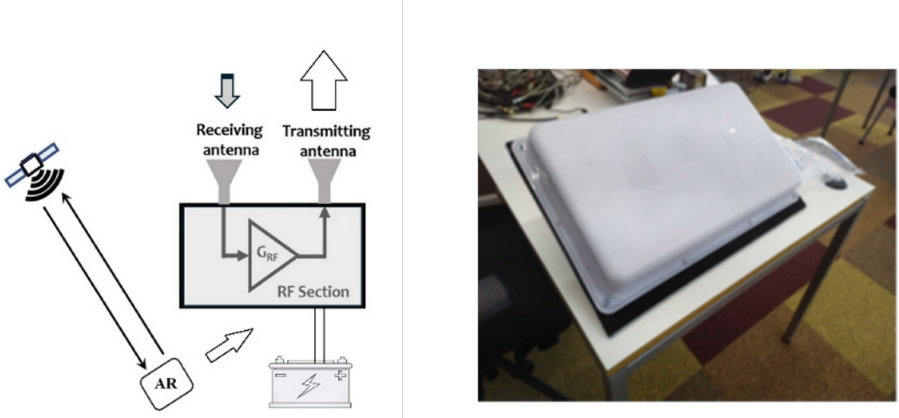


Fig. 2. Left: Working principle of the AR; right: Photo of a recent prototype available at CTTC.

An AR is available on the market since several years: the Electronic Corner Reflector (ECR-C) from Metasensing [11]; more recently a device designed to be used with the same satellite imagery, has been proposed in [12]; data discussed in this paper have been acquired using this device. The C-band AR proposed as an alternative to passive reflectors in a context where the required accuracy is of the order of a few millimeters as in application cases of interferometry used for deformation and land-slides monitoring with Sentinal-1 imagery. In the last decade some papers discussed the potential and issues of compact active transponders on an experimental basis [13–15]. The AR used in [12] was designed to provide a sufficiently stable phase response with reduced encumbrance and low power consumption and cost. In this paper we compare the accuracy achieved using this prototype to the observation's outcoming from some literature. We summarize the main outcomes achieved in [16] in the following section before discussing the comparison between AR and PCR.

3 A Test Based on Experimental Data

The data concerning this analysis was collected with the aim of monitoring an area threatened by a rockslide landslide. In a previous study [10], the AR was used as a reference point for the calculation of the interferometric phase of a set of PCRs, all

distributed in a large unstable area to provide a set of PSs for monitoring a large area. The orientation of the AR is azimuth North-100° and elevation East-51°, its position: Lat. 42° 26' 53" N, Long. 1º 29' 00" E, altitude: 980 m a.s.l. The selected area is indeed an area with very unstable temporal coherence due to the presence of vegetation. The AR is powered by a battery (12 V and capacity 70 Ah) and a solar panel (60 W). In order to evaluate the deformation data, 114 SAR images from Sentinel-1 SLC, descending geometry, orbit 110 and VV polarization were used. The interval covers the period from 4 December 2020 to 3 August 2023 for a total of 413 interferograms. The data was processed using the CTTC Persistent Scatterer Interferometry (PSI) processing chain [17]. The Digital Elevation Model (DEM) available from the Shuttle Topographic Mission (SRTM) was used with a resolution of 30 m. For interferometric processing, acquisitions in IW mode were processed. Figure 3a and Fig. 3b show two amplitude images obtained after applying a coherent average over 70 images, when the AR was not present, and 114 images with the AR working respectively. Both images include the response of a square trihedral corner reflector whose inner edge length is 0.60 m always working. The AR response clearly outstands with respect to that of PCR. Before analyzing the phase data, it is important to verify also if the amplitude values maintain sufficiently high and stable as previously commented. In Fig. 4 we show the amplitude (backscattering coefficient) obtained from the GRID images vs the acquisition date of the S1 images; except for some data missing due to battery drawbacks the average and standard deviation of the three data sets are satisfactory stable (see Table 1):

Table 1. Mean and standard deviation (STD) of the backscattering coefficient series for the AR and the surrounding clutter.

Target	Mean(dB)	STD(dB)	SCR(dB)
AR	+11.23	0.838	+24.00
Clutter	−21.17	1.803	+32.40

The corresponding trend for the phase data is shown in Fig. 5 which shows the data retrieved from the interferometric processing vs. the acquisition date. As expected, we observe the presence of periodic behavior which can be correlated to the seasonal variation of the air temperature at the satellite pass time on the AR electronics; anyway this correlation appear weak, as confirmed through the analysis of temperature internal to the case [16], but this can be ascribed to the limited range of variation of temperature values. Table 2 reports the mean, and the standard deviation of the displacement values plotted in Fig. 5.

Over the whole period of observation, the AR maintains the fluctuations within a two millimeters range of variation and 0.5 mm of bias. In previous paper focused on the use of a network of PCR during the same campaign [10] the stability associated to a stable PCR was estimated a bit worse than 1 mm per year; of note that this data is obtained over lapse of time, one year, shorter than that of the data of AR here analyzed which cover almost three years. Considering the accuracy required during a landslide or deformation monitoring in Line of Sigh (LOS) direction, this value can be considered satisfactory.

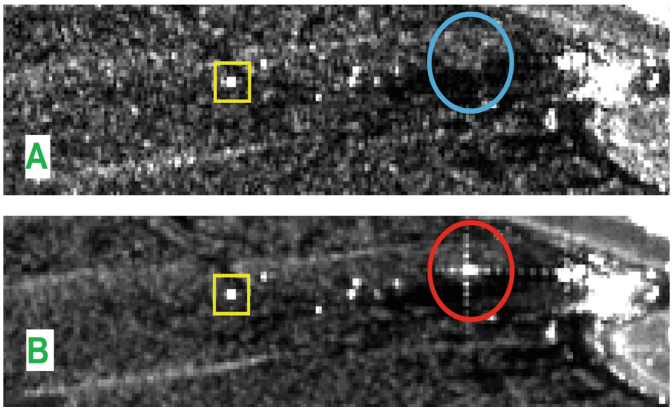


Fig. 3. Example of mean amplitude images with and without the AR. The yellow square indicates the PCR response. (After [16]); A without the AR, B with the AR switched on.

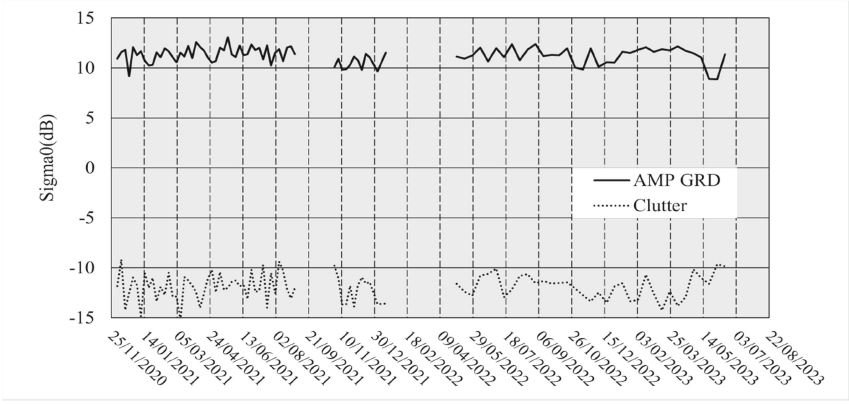


Fig. 4. Temporal series of amplitude for the AR.

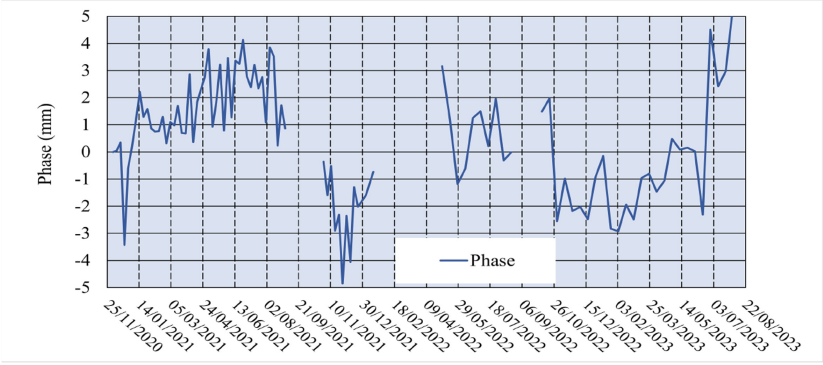


Fig. 5. Temporal series of the differential phase for the AR.

Table 2. Main statistical parameters of the phase data calculated form the temporal series expressed in mm of LOS displacement.

Target	Mean	STD	Min	Max
AR Phase (mm)	0.50	2.11	−5	6

4 Conclusions and Remarks

The values recalled in the previous section must be compared with data obtained from other experimental campaigns. One of the first comparisons of active reflectors' performance is in [13], where the applicability of active radar transponders for deformation measurements is studied and validated. Using more than one hundred of transponder-InSAR observations, from Envisat and ERS-2 imagery, and independent validation measurements such as levelling and the GPS, the authors achieve an empirical precision range of 1.8–4.6 mm, for double-difference (spatial and temporal) transponder phase measurements in the radar LOS. The study includes different scenario conditions as a quiet meadow and in an operational field case. On this basis, the authors consider a transponder "a compact and lightweight alternative to a corner reflector" [13]. This value is close to that achieved in [16] nevertheless the different device and incidentally data source (Sentinel-1). In a subsequent paper based on the use of the same device and using Sentinel-1, a longer set of time series, corresponding to almost one thousand measurements over 21 months is analyzed [14]. They found a partial correlation with surface temperature changes and the phase stability compared to a stable reference passive reflector corner is estimated to be 0.5–1.2 mm. The capacity to provide a stable reference also in geodetic survey is analyzed; the phase is confirmed to be significantly correlated with environmental temperature variations, showing variations within a range of 6 mm. This aspect demands a specific calibration for each device, as suggested in [12]. It depends on the specific case study, and this demands a careful analysis of the clutter signal characteristics. A most recent paper focuses on the usability and long-term performance of the AR referred as Compact Active Transponder (CAT) evaluated using two CATs from METASENSING in a field test [15]; in the study a vertical displacement of a CAT has been achieved using InSAR with an accuracy better than 1 cm, possibly a few millimeters and that the mean vertical velocity calculated from 14 months of InSAR displacement time series, can be determined with an accuracy of a few mm/year. It is confirmed that the LOS phase error is generally found to be moderately correlated with temperature, and the LOS error is ruled by linear relationship between and temperature ranging between approx. 0.1 and 0.2 mm/°C, a value close to that achieved in [12]. Finally, the main recommendations from the lesson learn according to the few papers available on the operative use of ARs can be resumed in the following. The device must be characterized not only through standard laboratory tests but also by providing calibration reports carried out in field conditions or disposing of some passive reflectors close to their installation site. In case of geodetic use, the delay added by the device, which appears to be significantly different in different devices, see [12] and [15], must

be carefully considered. Finally, we underline that although at C band the main advantages are not strictly from a quantitative performance point of view, i.e. better stability or higher accuracy, the AR demands an easier and lower costs of installation especially in hard environments. These advantages become more effective for lower frequency radar observations, see for example recent new space-borne SAR missions as NISAR and Biomass, where the installation of PCR with sufficient RCS should require unfeasible size.

Acknowledgement. This work has been partially funded by AGAUR, Generalitat de Catalunya, through the Consoli-dated Research Group RSE, "Geomatics" (Ref: 2021-SGR-00536) and it was conducted as part of the EU-funded project PERMAPYRENEES (INTERREG-POCTEFA 2021–2027, EFA063/01).

References

1. Nopphawan, T., Ichikawa, D., Katiyar, V., Nagai, M., Nagai, Y., Eguchi, T.: The evaluation of corner reflectors performance for utilization of Sentinel-1 data. J. Evol. Space Activities **1**, 71 (2023). https://doi.org/10.57350/jesa.71
2. Marinkovic, P., Ketelaar, G., van Leijen, F., Hanssen, R.: InSAR quality control-analysis of five years of corner reflector time series. In: Proceedings of Fringe Workshop (ESA SP-649), pp. 26–30, Frascati, Italy, November 2007
3. Xia, Y., Kaufmann, H., Guo, X.: Differential SAR interferometry using corner reflectors. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium, vol. 2, pp. 1243–1246, June 2002
4. Azadnejad, S. et al.: A low-cost passive corner reflector design for SAR and InSAR applications. Remote Sens. Environ. **338**, 115372 (2026). https://doi.org/10.1016/j.rse.2026.115372
5. Ulaby, F., Moore, R., Fung, A.: Microwave Remote Sensing Active and Passive, vol. 3. Artech House Inc., Norwood, MA (1986)
6. Crosetto, M., Monserrat, O., Cuevas-González, M., Devanthéry, N., Crippa, B.: Persistent scatterer interferometry: a review. ISPRS J. Photogramm. Remote Sens. **115**, 78–89 (2016)
7. Czikhardt, R., van der Marel, H., van Leijen, F.J., Hanssen, R.F.: Estimating signal-to-clutter ratio of InSAR corner reflectors from SAR time series. IEEE Geosci. Remote Sens. Lett. **19**, 1–5 (2022). https://doi.org/10.1109/LGRS.2021.3070045
8. Doerry, A.W., Brock, B.C.: Radar cross section of triangular trihedral reflector with extended bottom plate. In: Sandia Nat. Laboratories, Albuquerque, NM, USA, Tech. Rep. SAND2009–2993 (2009)
9. Garthwaite, M.: On the design of radar corner reflectors for deformation monitoring in multi-frequency InSAR. Remote Sens. **9**(7), 648 (2017). https://doi.org/10.3390/rs9070648
10. Luzi, G. et al.: A low-cost active reflector and a passive corner reflector network for assisting landslide monitoring using multitemporal InSAR. Remote Sens. Lett. **13**(11), 1080–1089 (2022). https://doi.org/10.1080/2150704x.2022.2122891
11. METASENSING: Radar Solutions (2025). https://metasensing.com/wp-content/uploads/2022/07/MetaSensing-ecr-c.pdf
12. Luzi, G., Espín-López, P.F., Mira Pérez, F., Monserrat, O., Crosetto, M.: A low-cost active reflector for interferometric monitoring based on Sentinel-1 SAR images. Sensors. **21**(6), 2008 (2021). https://doi.org/10.3390/s21062008

13. Mahapatra, P.S., Samiei-Esfahany, S., van der Marel, H., Hanssen, R.F.: On the use of transponders as coherent radar targets for SAR interferometry. IEEE Trans. Geosci. Remote Sens. **52**(3), 1869–1878 (2014). https://doi.org/10.1109/TGRS.2013.2255881
14. Czikhardt, R., van der Marel, H., Papco, J., Hanssen, R.F.: On the efficacy of compact radar transponders for InSAR geodesy: results of multiyear field tests. IEEE Trans. Geosci. Remote Sens. **60**, 5215913 (2022)
15. Meister, A., Balasis-Levinsen, J., Keller, K., Pedersen, M.R.V., Merryman Boncori, J.P., Jensen, M.: A field test of compact active transponders for InSAR geodesy. J. Geodetic Sci. **14**(1), 20220164 (2024). https://doi.org/10.1515/jogs-2022-0164
16. Luzi, G., Gao, Q., Espín-López, P.: Experimental study of the stability of a low-cost C-band active reflector using Sentinel-1 imagery. IEEE Access. **13**, 55202–55210 (2025). https://doi.org/10.1109/ACCESS.2025.3554305
17. Devanthéry, N., Crosetto, M., Monserrat, O., Cuevas-González, M., Crippa, B.: An approach to persistent scatterer interferometry. Remote Sens. **6**(7), 6662–6679 (2014). https://doi.org/10.3390/rs6076662



Statistical Anomaly Detection Using LC-Curves: An Application to Geophysical Time Series

Fuad Aleskerov[1,2] and Aleksei Kosychev[1(✉)]

[1] HSE University, Myasnitskaya St. 20, Moscow 101001, Russia
amkosychev@edu.hse.ru
[2] Institute of Control Science of Russian Academy of Science, Profsoyuznaya St. 65, Moscow 117997, Russia

Abstract. We adapt the Life Cycle (LC) curve methodology to the problem of statistical anomaly detection in time series. LC-curves transform a raw non-negative time series into a cumulative share profile that highlights when, within a fixed observation window, most of the signal intensity is concentrated. Deviations of this profile from the uniform-accumulation baseline serve as simple anomaly indicators that require no distributional assumptions. As an application, we apply the method to 5-min GPS displacement time series from the Nevada Geodetic Laboratory, comparing 196 earthquakes with $M \geq 6$ against 104 seismically quiet intervals, and evaluate performance with Probability of Detection (POD), False Alarm Ratio (FAR), and Critical Success Index (CSI). The method achieves POD ≈ 0.82 (CSI $= 0.62$, FAR $= 0.29$) at the lowest tested threshold. Pre-event anomalies are statistically distinguishable from calm-period anomalies, demonstrating that LC-curves can effectively convert high-frequency time series into actionable anomaly indicators.

Keywords: LC-curves · non-stationary time series · anomaly detection · non-parametric methods · event-based verification metrics

1 Introduction

Short-term earthquake prediction remains a difficult and controversial problem, despite decades of work on seismic and geodetic precursors. GPS station networks now provide small-scale measurements of ground motion at high temporal resolution, offering the possibility to detect subtle deformation anomalies in the hours to days before large events. However, it is still unclear to what extent such signals can be distinguished from normal variability in a systematic, statistically robust way, and how often similar anomalies appear during seismically quiet periods.

In this study, we address this question by adapting the Life Cycle (LC) curve methodology — originally developed for analyzing non-stationary customer behavior — to the problem of pre-earthquake ground-deformation anomalies. LC-curves method transforms a raw time series into a cumulative profile that highlights when, within a fixed window, most of the signal is accumulated. Applied to 5-min horizontal GPS displacement

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 384–392, 2027.
https://doi.org/10.1007/978-3-032-30539-8_30

series from the Nevada Geodetic Laboratory, this representation allows us to compress pre-event windows into a small number of anomaly indicators.

We build a global dataset that treats earthquakes and calm intervals symmetrically. For 196 earthquakes with $M \geq 6$ and sufficient geodetic coverage, as well as for 104 selected seismically quiet intervals, we compute LC-curve-based anomaly measures from nearby GPS stations. Each event (or calm case) is then represented by a single maximum anomaly value, enabling event-based evaluation with standard verification metrics: Probability of Detection (POD), False Alarm Ratio (FAR), and Critical Success Index (CSI). This framework allows us to quantify both apparent precursors and false alarms under consistent assumptions.

2 Literature Review

Research on earthquake prediction spans more than a century, yet classical seismological approaches have long struggled to deliver reliable short-term predictions. As Geller [8] argued, precursor signatures such as seismic quiescence, geomagnetic changes, or hydrological anomalies often lack reproducibility across events and regions. Later studies refined the understanding of crustal deformation [6, 15], while comprehensive reviews summarized decades of precursor research [10]. Together, these works highlight both the diversity of potential precursors and the difficulty of converting them into operational forecasts.

The emergence of artificial intelligence marked a major shift to data-driven modeling. Early reviews [3, 4] showed that machine learning methods (SVM, neural networks, fuzzy–neural hybrids) outperformed classical statistical techniques, especially for nonlinear and noisy datasets. Subsequent regional studies [7, 11] reinforced the value of ML for forecasting ground-motion parameters and short-term seismic behavior in local networks.

Deep learning further transformed earthquake data analysis. The EQTransformer model [12] demonstrated high accuracy in waveform detection and P/S phase picking, enabling the construction of high-resolution catalogs with up to ten times more events than traditional methods. Other works [5] proposed that such dense catalogs reveal previously hidden triggering patterns and fault interactions.

Multi-sensor approaches extend these capabilities. GPS and InSAR displacement time series [9] and satellite-derived thermal and atmospheric indicators [17] have shown measurable anomalies days to weeks before major events. Meanwhile, IoT-enhanced systems [6, 14] integrate ML with dense, real-time sensor networks, improving latency and scalability for early-warning applications.

Overall, the literature shows a clear progression from precursor-centric approaches with limited reliability toward ML- and DL-based methods that extract predictive structure directly from seismic, geodetic, and satellite data, increasingly combined with multi-sensor fusion. Despite remaining challenges (data scarcity for large earthquakes, interpretability, regional variability), current advances provide strong evidence that high-resolution, multimodal data combined with modern analytical approaches can reveal meaningful short-term predictive patterns previously inaccessible to traditional methods.

3 Life Cycle (LC) Curves

The LC-curves approach provides a way to analyze the inherently non-stationary behavior of a system by examining how a chosen indicator evolves over a given time horizon. Unlike traditional econometric techniques for non-stationary time series where the temporal dimension typically enters the model as a parameter, the LC-curves method focuses directly on the cumulative dynamics of the process. This framework makes it possible to treat both stationary and non-stationary time series without requiring any assumptions about stationarity.

Conceptually, an LC-curve represents the life cycle of an object (e.g., a customer in the original formulation). The foundational models behind this methodology were first introduced in [1] and further developed in [2].

Earlier applications of the methodology to earthquake data were first presented in a plenary talk at the eighth International Conference on Dynamics of Disasters (DOD 2025), Nanjing, China, 2025 (forthcoming [19]), and at the 13th International Conference on Geological and Civil Engineering (ICGCE 2026), Fukuoka, Japan, 2026.

3.1 Construction of LC-Curves

Let n be the number of time moments, and for a particular customer non-negative time series is defined as $\{Vi\}$, $Vi \geq 0$, $i = 1, \ldots, n$. To construct an LC-curve at time t the sum of the values of the time series Vi is calculated starting from time t, till time n. Then the share of this sum to the sum of all values of Vi from 1 to n is calculated. Namely,

$$LC(t) = \frac{\sum_{i=t}^{n} v_i}{\sum_{i=1}^{n} v_i}$$

Where Vi is the value of the time series at time i.

Notice that in the first moment of time $t = 1$. $LC(1) = 1$. At the last time moment $t = n$, $LC(n)$ is not always close to 0, for instance, if Vi is equal to 0 for all i, except n. then $LC(n) = 1$. Nevertheless, for the most of the cases, $LC(n)$ is close to 0 as the number of periods of time n increases.

3.2 Examples

Let us consider two examples taken directly from the GPS dataset. We use the vertical displacement record at station CONZ, located 40 km from the epicenter of the M = 6.2 earthquake (event usp000hkg9). On a seismically quiet day at this station the input series consists of small fluctuations, and the corresponding LC-curve almost coincides with the uniform-accumulation bisector (Fig. 1).

In the second example (Fig. 2) we use the same station and component on the day of the earthquake. The input series is dominated by a sharp burst of displacement at the moment the seismic wave reaches the station. The LC-curve responds with a pronounced step and stays well below the bisector for the remainder of the window.

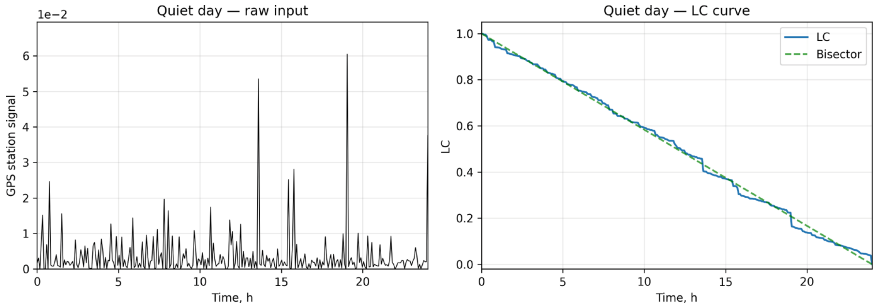


Fig. 1. A GPS displacement series at station CONZ on a seismically quiet day

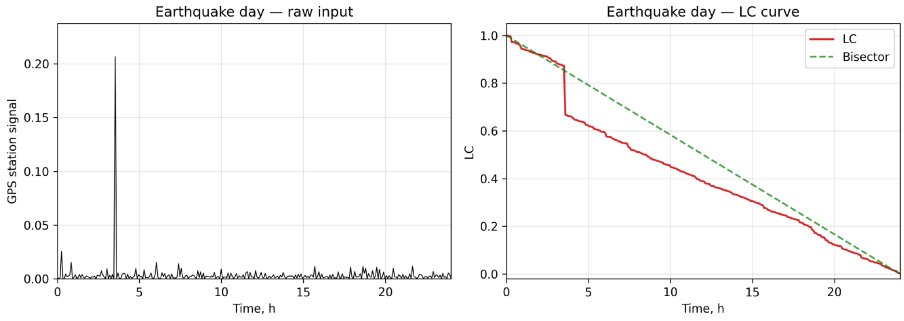


Fig. 2. GPS displacement series at station CONZ before earthquake.

4 Data Description

This study relies on several datasets that together enable a systematic analysis of GPS anomalies preceding major seismic events. We begin with a global catalogue of earthquakes with magnitude $M \geq 6$, containing 2750 events in the 2004–2025 period (U.S. Geological Survey [16]), and compute the distance between each event and every GPS station available from the Nevada Geodetic Laboratory (NGL) [13]. We then restrict the dataset to earthquakes that have at least one GPS station within 50 km radius from the epicenter, resulting in a total of 590 events.

To ensure that the selected earthquakes are not influenced by earlier local seismicity, we incorporate a second global catalogue that includes all earthquakes with $M \geq 2$ covering the same time period. For each of the 590 earthquakes, we analyze the surrounding seismic activity in a pre-event window defined by 14 days prior to the mainshock and within a 30 km radius. Events that are not preceded by any smaller earthquakes in this window form a filtered subset of 329 earthquakes with $M \geq 6$.

We complement these seismic datasets with a global catalogue of 22 287 GPS stations. For each of the 329 selected earthquakes, we identify all stations located within 50 km radius, forming 1919 earthquake–station pairs. From these pairs, we successfully download 1161 unique GPS time series from NGL. Each downloaded file contains 5-min observations of the north–south, east–west, and vertical displacement components

recorded throughout a year. Out of the 1161 files, 978 contain a 4-day interval preceding the earthquake and can therefore be processed by the model. The processed files correspond to 196 unique earthquakes.

To evaluate false alarms, we additionally analyze GPS displacements during seismically quiet periods. For each of the 176 stations associated with the 196 earthquakes, a calm window is defined as an interval between two successive $M \geq 2$ events located within 50 km of the station, drawn from the same global catalogue used in the earthquake-screening step. Only calm windows of at least 124 days are retained — this margin ensures that the 4-day analysis segment lies well inside a seismically quiet interval — and for every station we select the longest such window in which a continuous 4-day block of 5-min observations is available. This procedure yields 104 fully complete calm-period datasets, each suitable for LC-curve computation in the same 4-day format used for the earthquake cases.

5 Evaluation Metrics

To assess the performance of the model in detecting earthquakes, we employ three widely used verification metrics Probability of Detection (POD), False Alarm Ratio (FAR), and Critical Success Index (CSI). These metrics are standard in event-based evaluation, particularly in applications involving rare or spatially sparse phenomena [18].

In all metric definitions, we use the standard contingency-table notation, where TP (true positives) denotes correctly detected events, FP (false positives) represents cases where an event was predicted but did not occur, and FN (false negatives) corresponds to events that occurred but were not predicted.

POD measures the model's ability to correctly identify positive events.

$$POD = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN})$$

A POD value of 1 indicates perfect detection.

FAR describes the fraction of predicted events that did not actually occur.

$$FAR = \mathrm{FP}/(\mathrm{TP} + \mathrm{FP})$$

Lower FAR values indicate fewer false alarms.

CSI provides an overall measure of event-detection performance by jointly penalizing both missed events and false alarms.

$$CSI = \mathrm{TP}/(\mathrm{TP} + \mathrm{FP} + \mathrm{FN})$$

Unlike many other metrics, CSI does not include true negatives, making it particularly suitable for sparse-event domains.

6 LC-Curve Model Implementation

In this section we describe the LC-curve model implementation for earthquake cases and calm-day scenarios to verify the true nature of anomalies in earth displacements.

First, we take the nearest GPS stations to the earthquake. Then, from the next point in the station's time series (i) we subtract the value of the previous one (i − 1). To get the correct results when using the LC-curves method, we need to get rid of the negative values of the earth's surface displacements. To do this, we extract the minimum values for each of the displacement directions from the data obtained. The resulting values are then converted by raising to the fourth power. Lower powers and the absolute value were also tested. The fourth power gave the strongest separation between bursts and background noise and was therefore retained.

As a scalar deviation metric for each LC-curve we use its maximum Euclidean distance to the uniform-accumulation bisector. This single value is referred to throughout as the anomaly value and is the quantity to which all detection thresholds in the following sections are applied. In fact, much more extended metrics can be used as well.

6.1 LC-Curve Analysis of Pre-Seismic Displacements

The number of GPS stations available per earthquake varies substantially. According to the processed dataset, 87 earthquakes are associated with only one GPS station, while the remaining events involve multiple stations, in some cases extending to more than ten stations per event. This variability reflects differences in regional network density but does not affect the workflow, because we retain only the station exhibiting the largest LC-curve deviation as the observation for each earthquake. This procedure results in a single anomaly metric per event, enabling consistent event-level evaluation.

To quantify the model's ability to detect earthquake-related anomalies, we compute true-positive (TP) and false-negative (FN) counts for a sequence of anomaly thresholds ranging from 0.09 to 0.15. The resulting detection statistics are summarized in Table 1.

Table 1. True positives and false negatives for all earthquakes at different thresholds.

Threshold	TP	FN
0.09	161	35
0.10	145	51
0.11	125	71
0.12	109	87
0.13	103	93
0.14	88	108
0.15	82	114

6.2 Model Performance for Non-Earthquake Scenarios

To evaluate the baseline behavior of the LC-curves model in the absence of seismic activity, we apply the same anomaly-detection procedure to long, seismically quiet intervals.

Following the multi-stage filtering and completeness checks described in Sect. 4, we obtain 104 stations that meet the criteria for uninterrupted observations.

For each station, we compute the LC-curve deviation for the selected calm-day interval and extract the maximum anomaly value. We evaluate the number of calm-day anomalies that exceed a series of thresholds from 0.09 to 0.15. For each threshold, we count the number of false positives (FP) and true negatives (TN). The results are summarized in Table 2.

Table 2. False positives and true negatives for 104 calm-period intervals.

Threshold	FP	TN
0.09	65	39
0.10	50	54
0.11	39	65
0.12	32	72
0.13	22	82
0.14	19	85
0.15	14	90

At the lowest thresholds (0.09–0.10), the model registers 65 and 50 false positives, reflecting that a small fraction of calm-day intervals can occasionally exhibit moderate deviations. As the threshold increases, FP counts decrease steadily from 39 at 0.11 to 14 at 0.15, while TN counts increase correspondingly. This monotonic transition indicates that higher anomaly thresholds effectively suppress non-earthquake detections.

7 Discussion

In this section, we evaluate the model's ability to discriminate between earthquake-related anomalies and background variability by comparing all earthquake cases against all non-earthquake cases. We apply the same set of event-based metrics POD, FAR, and CSI across a sequence of detection thresholds. The resulting metrics are summarized in Table 3.

Table 3. Detection metrics (POD, FAR, CSI) for 196 earthquakes vs. 104 calm-period intervals.

Threshold	POD	FAR	CSI
0.09	0.82	0.29	0.62
0.10	0.74	0.26	0.59
0.11	0.64	0.24	0.53

(*continued*)

Table 3. (*continued*)

Threshold	POD	FAR	CSI
0.12	0.56	0.23	0.48
0.13	0.53	0.18	0.47
0.14	0.45	0.18	0.41
0.15	0.42	0.15	0.39

Across the entire threshold range, we observe a monotonic decline in POD, indicating that stricter detection requirements inevitably lead to an increase in missed events. At the same time, the FAR steadily decreases, reflecting a reduction in false alarms as the model becomes more conservative. The CSI also declines smoothly with increasing threshold. Since CSI penalizes both types of errors (misses and false alarms), the gradual reduction of CSI indicates an overall deterioration of balanced performance as the model is shifted into a low-sensitivity regime. Importantly, the maximum CSI ≈ 0.62 is observed at the lowest threshold, which corresponds to the highest POD of 0.82 but also the highest FAR of 0.29.

Taken together, these patterns provide a systematic view of the model's performance and allow us to identify threshold ranges where the balance between missed detections and false alarms is most appropriate for operational use.

8 Conclusion

This study examined whether LC-curves, originally developed to describe customer behavior, can be adapted to detect short-term GPS displacement anomalies preceding large earthquakes. Using LC-based anomaly measures for 196 earthquakes ($M \geq 6$) and 104 seismically quiet intervals, we evaluated detection skill with POD, FAR, and CSI across a range of anomaly thresholds. The results show that LC-curves do capture statistically distinct patterns in pre-seismic GPS time series compared to calm periods, with the highest POD ($\approx$0.82) at the lowest tested threshold of 0.09.

Overall, our findings suggest that LC-curves provide a simple way to compress high-frequency GPS time series into anomaly indicators that carry modest predictive value for large earthquakes.

Acknowledgments. This work has been supported by the grants of the Russian Science Foundation, RSF No. 24-61-00030, https://rscf.ru/project/24-61-00030/.

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

References

1. Aleskerov, F., Golubenko, M.: On the Evaluation of a Symmetry of Political Views and Polarization of Society. Working paper WP7/2003/04. HSE University, Moscow (2003)
2. Aleskerov, F., Volsky, V., Yakuba, V.: Application of the symmetry and polarization indices to the analysis of consumption dynamics of customers. In: Proceedings 4th International Conference on Control Problems, Moscow (2009)
3. Azam, F., Sharif, M., Yasmin, M., Mohsin, S.: Artificial intelligence based techniques for earthquake prediction: a review. Sci. Int. **26**(4), 1495–1502 (2014)
4. Banna, M.H.A., et al.: Application of artificial intelligence in predicting earthquakes: state-of-the-art and future challenges. IEEE Access. **8**, 192880–192923 (2020)
5. Beroza, G.C., et al.: Machine learning and earthquake forecasting — next steps. Nat. Commun. **12**, 4761 (2021)
6. Bilal, M.A., et al.: An early warning system for earthquake prediction from seismic data using batch normalized graph convolutional neural network. Sensors. **22**(17), 6482 (2022)
7. Essam, Y., et al.: Exploring the reliability of different artificial intelligence techniques in predicting earthquake for Malaysia. Soil Dyn. Earthq. Eng. **147**, 106826 (2021)
8. Geller, R.J.: Earthquake prediction: a critical review. Geophys. J. Int. **131**(3), 425–450 (1997)
9. Gitis, V., Derendyaev, A., Petrov, K.: Analyzing the performance of GPS data for earthquake prediction. Remote Sens. **13**(9), 1842 (2021)
10. Huang, F., et al.: Studies on earthquake precursors in China: a review for recent 50 years. Geod. Geodyn. **8**(1), 1–12 (2017)
11. Marhain, S., et al.: Investigating the application of artificial intelligence for earthquake prediction in Terengganu. Nat. Hazards. **108**(2), 977–999 (2021)
12. Mousavi, S.M., et al.: Earthquake transformer — an attentive deep-learning model for simultaneous earthquake detection and phase picking. Nat. Commun. **11**, 3952 (2020)
13. Blewitt, G., Hammond, W.C., Kreemer, C.: Harnessing the GPS data explosion for interdisciplinary science. Eos. **99** (2018)
14. Pwavodi, J., et al.: The role of artificial intelligence and IoT in prediction of earthquakes: review. Artif. Intell. Geosci. **5**, 100075 (2024)
15. Shen, Z., Jackson, D.D., Ge, B.X.: Crustal deformation across and beyond the Los Angeles Basin from geodetic measurements. J. Geophys. Res. **101**, 27957–27980 (1996)
16. Geological Survey, U.S.: Earthquake Hazards Program: Lists, Maps, and Statistics. https://www.usgs.gov/programs/earthquake-hazards
17. Xiong, P., et al.: Towards advancing the earthquake forecasting by machine learning of satellite data. Sci. Total Environ. **771**, 145256 (2021)
18. Wilks, D.S.: Statistical Methods in the Atmospheric Sciences, 3rd edn. Academic (2011)
19. Aleskerov, F., Kosychev, A.: Observations of ground surface displacement as a factor in earthquake onset using the LC-curves method. In: Dynamics of Disasters, Springer Optimization and its Applications, vol. 229. Springer (2026)

A Robotic GPR Pipeline for Automated Subsurface Analysis in Multiple Scenarios

Nuria Rodríguez Calderón[1(✉)], Berta Viñes Santamaria[1], David Caballero Flores[1], Tahzeeb Hussain[1], and Vega Perez-Gracia[2]

[1] Centre de Disseny d'Equips Industrials (CDEI), Universitat Politècnica de Catalunya, Barcelona, Spain
nuria.rodriguez.calderon@upc.edu

[2] Universitat Politècnica de Catalunya-BarcelonaTech. Dpt. Resistència de Materials i Estructures a l'Enginyeria. Escola d'Enginyeria de Barcelona Est., Barcelona, Spain

Abstract. GPR is a technique for shallow subsurface sensing, but its use is limited by variable acquisition conditions and labor-intensive interpretation. This work presents a robotic GPR pipeline for agricultural and construction scenarios, combining two complementary sensing platforms with automated radargram processing. The framework includes radargram standardization, AI-based hyperbola detection, geometric fitting, circular slice generation, and frequency-domain analysis. Results show root density related patterns, as well as discrimination between metallic and plastic pipes. Additional branches for soil-moisture estimation and material classification are under development.

Keywords: Ground-penetrating radar (GPR) · robotic surveying · hyperbola detection · subsurface analysis

1 Introduction

Ground-penetrating radar (GPR) is a non-destructive electromagnetic technique for shallow subsurface investigation, as it enables the detection of buried structures and material interfaces through contrasts in dielectric properties [6]. It operates by emitting short pulses into the material and recording the reflections generated when the waves encounter discontinuities. These reflections are registered over time to form radargrams, in which two main feature types are typically observed: hyperbolic signatures, associated with localized buried objects detected as the antenna passes over them (e.g., reinforcement bars, pipes, or voids), and horizontal reflections, which indicate layer boundaries or changes in material composition.

Despite its potential, practical GPR use remains limited by operator-dependent acquisition and labor-intensive interpretation. Survey quality is sensitive to acquisition conditions, while radargrams often require expert preprocessing and manual inspection. To address both issues, this work proposes a

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 393–404, 2027.
https://doi.org/10.1007/978-3-032-30539-8_31

robotic GPR framework that combines repeatable acquisition with automated radargram processing for agricultural and construction scenarios.

2 Related Work

GPR has been extensively studied for subsurface characterization, with applications spanning the detection of underground utilities [18], assessment of infrastructure condition [14], soil mapping and moisture estimation [12]. However, the effectiveness of these applications strongly depends on the consistency of data acquisition. To improve the repeatability and scalability of GPR surveys, recent research has focused on integrating GPR sensors into robotic platforms [8]. Most robotic GPR systems rely on ground mobile platforms to ensure stable antennaground coupling. Other configurations have been also explored: crawler-based where manual access is difficult [3], in-pipe robots for internal inspection of underground structures, and UAV-based robotic systems [11]. Despite these advances, robotic GPR systems remain limited by localization challenges (especially in GNSS-denied environments) and difficulties in maintaining consistent acquisition conditions.

In parallel, to address radargram interpretation challenges, research has focused on automated and semi-automated methods based on signal processing and machine learning. More recently, deep learning methods have demonstrated strong performance in tasks such as hyperbola detection and GPR denoising [1], enabling more scalable interpretation pipelines. Nevertheless, their performance remains limited by data scarcity and domain generalization issues [17].

However, these two research directions (robotic data acquisition and automated interpretation) have largely evolved independently, highlighting the need for integrated frameworks that jointly address both aspects.

3 Robotic Platforms and GPR Sensing System

The overall approach is conceived as a GPR-centred system in which robotic platforms act as acquisition enablers and the downstream processing chain transforms radar measurements into application-relevant information. To accommodate different operational requirements, the framework relies on two complementary robotic platforms (Fig. 1). AGROMOBY [5] is designed for agricultural and field environments, where large-area coverage, uneven terrain, and stable ground-following acquisition are essential. The platform is coupled with MALÅ impulse GPR antennas with single-frequency configurations ranging from 200 to 2500 MHz, allowing the operating frequency to be selected according to the required trade-off between penetration depth and spatial resolution. The antenna is mounted on an external non-conductive support with a passive wheel to maintain a near-constant antenna-ground distance.

In contrast, OLIWALL [7] is conceived for demolition scenarios, where sensing must cover horizontal and elevated structures. The platform integrates a compact multi-frequency GPR system, the Screening Eagle Proceq GP8800, suited to

short-range, very-high-resolution inspection tasks such as reinforcement bars or internal structural features. The sensor is mounted on a robotic manipulator combined with an elevation column, enabling inspection of overhead surfaces up to approximately 2.8 m.

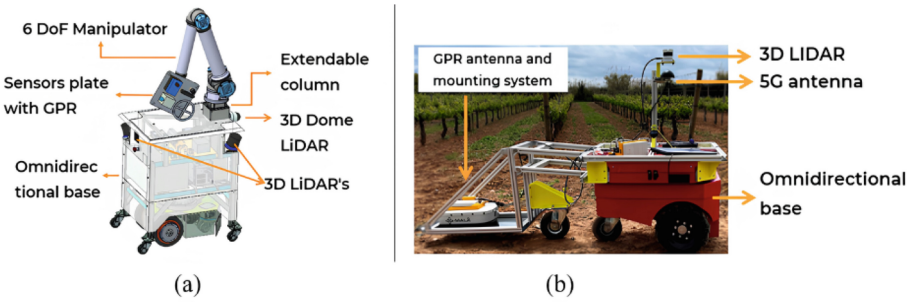


Fig. 1. Robotic platforms for GPR sensing: (a) OLIWALL, a mobile manipulator with arm-mounted GPR for structural inspection; (b) AGROMOBY, an agricultural robot with ground-supported GPR for field surveys.

From a motion-control perspective, both platforms ensure repeatable and structured scanning trajectories. The current acquisition strategy does not rely on a conventional wheel-based GPR odometer; instead, trace positioning combines time-based acquisition, electronic odometry, and robot localization under approximately constant-speed motion. In AGROMOBY, ROS2-based navigation fuses LiDAR, GNSS, IMU, and vision for autonomous field traversal under consistent sampling conditions, with the antenna operated in a near-contact configuration and an air gap kept below approximately 20 cm; larger stand-off distances reduce antenna coupling. In OLIWALL, the target configuration is contact acquisition against the inspected surface, and a non-wheel, contact-based mechanical odometer is under design. In both cases, robotic–GPR integration is handled through ROS2 interfaces; for the Proceq GP8800, communication additionally relies on the manufacturer's API.

Despite these capabilities, several practical challenges remain. Mechanical vibrations can introduce radar noise, maintaining a consistent antenna–surface distance is non-trivial, localization errors directly affect spatial interpretation, and electromagnetic interference from onboard electronics must be minimized. These issues highlight the importance of co-design between the robotic platform and the sensing system to ensure reliable GPR acquisition. The final outputs of the framework, as well as the specific interpretation pathway used to obtain them, depend on the target application. In construction scenarios, the system supports the detection and characterization of hidden structural elements and material contrasts, whereas in agricultural contexts it is oriented toward root-related analysis and soil-moisture-related variability.

4 GPR Data Processing and Interpretation

Once acquired, the data is converted into radargrams, which form the common input to the interpretation pipeline. As shown in Fig. 2, this pipeline combines a shared backbone with application-specific branches.

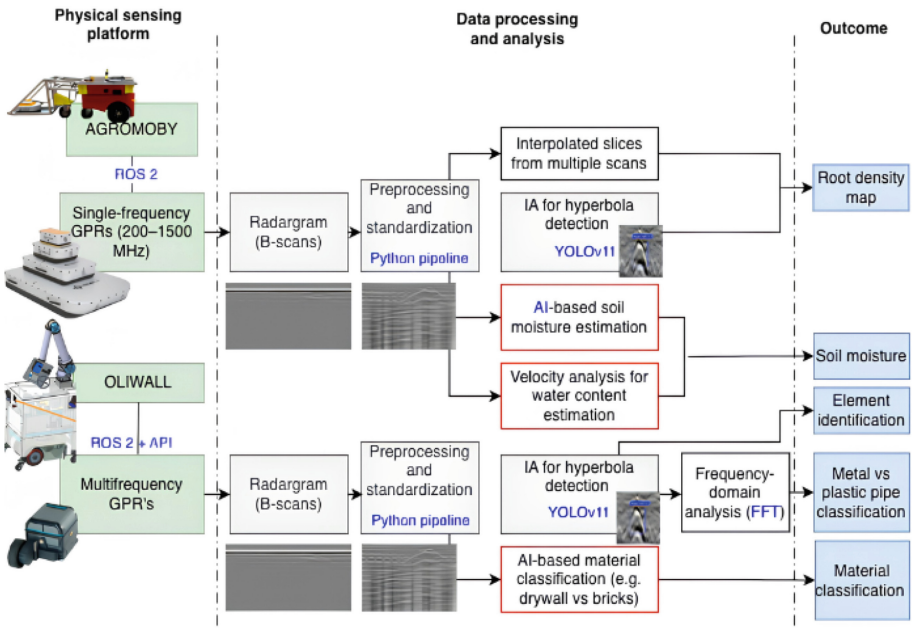


Fig. 2. Robotic GPR framework. Red: work in progress; grey: shared stages. (Color figure online)

4.1 Radargram Preprocessing and Standardization

Radargram standardization is essential for automated GPR interpretation. In conventional practice, visualization and preprocessing are often adjusted manually in commercial software such as `Reflex` or `GPR-SLICE`, which is not compatible with a robotic pipeline. In the present work, radargrams are standardized using basic metadata, including the number of traces and samples, from which an aspect ratio is defined automatically. Although formally a visualization choice, this becomes an important step when radargrams are used as inputs for automated analysis, particularly in image-based AI pipelines. A shared preprocessing sequence is then applied to all radargrams within the same application-specific outcome, based on prior expert knowledge. The current implementation uses automatic time-zero correction, background removal, dewow, low-frequency suppression, and linear gain within a Python workflow integrated into the ROS-based processing pipeline.

4.2 AI-Based Detection of Subsurface Hyperbolas

Deep learning techniques have been increasingly adopted for the automated detection of subsurface objects from GPR data in engineering and environmental applications [17]. As discussed previously, GPR radargrams typically contain two main types of signatures: hyperbolas, associated with localized buried objects, and horizontal reflections, associated with layer boundaries. Most automated interpretation efforts have focused on hyperbola detection, evolving from classical methods such as the Hough transform and template matching [19] toward convolutional neural networks and one-stage object detectors [2]. Among these, the YOLO family has shown strong performance due to its efficiency and accuracy [20], and has already been applied successfully to hyperbola detection and segmentation in GPR B-scans [4,15]. Following this line, the present work adopts YOLOv11M for hyperbola detection. In contrast, the automated interpretation of horizontal reflections remains less mature.

Training a reliable deep learning model requires large annotated datasets, yet acquiring and labelling GPR data is costly and expert-dependent. To mitigate this, the present work uses the publicly available GPR Scans Pipeline dataset from Roboflow Universe [10], containing annotated B-scan images for underground pipeline detection with bounding-box labels of hyperbolic signatures. This dataset is combined with experimental data acquired using the OLIWALL and AGROMOBY platforms. The dataset comprises 1,508 labelled B-scan images, split into training and validation subsets in an 80:20 ratio. YOLOv11M was fine-tuned from pre-trained weights (`yolo11m.pt`) for up to 100 epochs with early stopping and a batch size of 4. The best checkpoint was obtained at epoch 53, and the corresponding validation results are summarized in Table 1.

The balance between precision and recall suggests no pronounced bias toward false positives or negatives. Figure 3 shows representative validation detections. However, transferability to out-of-distribution data remains limited, with lower confidence and more false negatives in new environments. This reflects dataset bias and the current single-class formulation. Improving generalization will require larger, more diverse real datasets, potentially complemented with synthetic data generated with gprMax [21].

Table 1. Validation performance of the fine-tuned YOLOv11M model.

Metric	Value
mAP@50	0.878
mAP@50–95	0.439
Precision	0.855
Recall	0.852

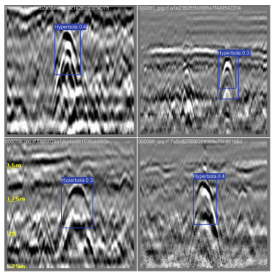


Fig. 3. Hyperbola detections.

4.3 Hyperbola Geometric Fitting: Estimation of Position, Depth and Apparent Size

The output of the AI-based detection stage consists of bounding boxes that localize hyperbolic signatures within the radargram. Since a detection model is used rather than a segmentation approach, these regions of interest do not directly describe the hyperbola geometry. Therefore, the detected regions are mapped back from the image representation to the original numerical GPR data, where each B-scan is treated as an amplitude matrix. Within each region, candidate points are identified as local amplitude maxima and evaluated according to geometric consistency criteria to determine whether they follow a hyperbolic pattern. For the valid points, a hyperbolic fitting procedure is applied using the analytical diffraction model:

$$t(x) = \sqrt{t_0^2 + \frac{4(x - x_0)^2}{v^2}} \tag{1}$$

where x_0 is the horizontal position of the target, t_0 is the two-way travel time at the apex, and v is the propagation velocity in the medium. Once fitted, the hyperbola apex is used to estimate the target position and depth d:

$$d = \frac{v \cdot t_0}{2} \tag{2}$$

The curvature and lateral extent of the fitted curve also provide an approximation of the apparent object size. This estimation requires knowledge, or at least reasonable assumption, of the propagation velocity in the medium. Although not directly known, it can often be approximated from contextual information. In construction scenarios, common materials such as concrete tend to exhibit similar velocity ranges, and OLIWALL provides complementary sensing information that helps select suitable values.

4.4 Frequency-Domain Analysis

This analysis corresponds to an application-specific processing branch aimed at distinguishing between metallic and plastic pipes. While recent studies have explored deep learning and feature-based approaches for material discrimination in GPR data [13], we adopt a simpler initial approach based on a physically interpretable descriptor: the spectral energy of the signal.

Once the hyperbolic signature has been localized through AI-based detection and geometric fitting, the corresponding GPR trace is transformed into the frequency domain using the fast Fourier transform (FFT), and the spectral energy is computed from the FFT magnitude as a comparative feature. The results shown in Fig. 4 are obtained from real experimental data acquired over intentionally buried pipes. Complementary analyses using simulated data generated with gprMax [21] lead to consistent observations. Metallic objects produce significantly higher spectral energy values than plastic ones, due to their higher

conductivity and stronger impedance contrast with the surrounding medium. As shown in Fig. 4, steel and aluminium exhibit clearly higher energy distributions than plastic, indicating that spectral energy is an effective feature for distinguishing metallic from non-metallic targets.

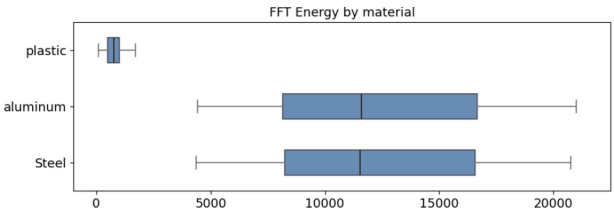


Fig. 4. Distribution of FFT-based spectral energy for different buried pipe materials, from real experimental data.

4.5 Slice-Based Spatial Analysis

Beyond point-wise analysis of individual radargrams, spatial interpretation can be enhanced through the generation of slices, which represent the distribution of signal amplitude at a given depth (or time sample) across multiple scans. These slices provide a planar view of the subsurface response and are particularly useful for identifying spatial patterns and structures that are not evident in single B-scans. In this work, slices are generated directly from the preprocessed GPR data using custom Python tools. While conventional slice generation typically assumes linear survey trajectories, a key contribution of this work is the extension to circular slice representations. This is particularly relevant for acquisition strategies based on circular scanning patterns, where the robot collects data along concentric paths. This choice is motivated by the approximately radial growth of tree roots and has been useful in our root-system analysis (see Sect. 5); similar concentric circular GPR scanning schemes have been used for root-system reconstruction [9].

4.6 Work in Progress Pipelines and Scenarios: Soil Moisture and Material Classification

In addition to the core processing backbone, several application-specific branches are still under development (highlighted in red in Fig. 2) and are therefore presented here as ongoing work rather than consolidated contributions.

One of these lines addresses soil-moisture estimation from GPR data. Since water content affects electromagnetic propagation velocity, classical methods typically require calibration with reflectors or objects buried at known depths [12], making them poorly suited to autonomous robotic operation. Preliminary field tests have therefore been combined with synthetic datasets generated with gprMax [21] to investigate feature correlations with soil water content [22].

Figure 5 presents peak instantaneous amplitude (PIA) as a function of water content, showing linear trends. In parallel, learning-based approaches are being explored, building on previous work by the authors [16], although these models have not yet been integrated into the full pipeline.

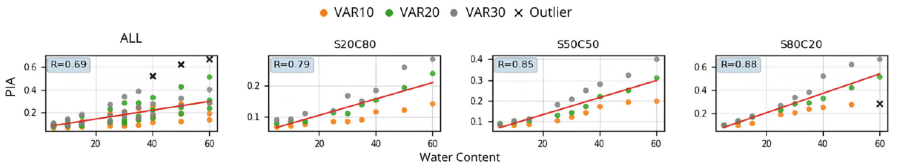


Fig. 5. Example of GPR a feature response (PIA) for 3 heterogeneous synthetic soils. Columns correspond to different sand–clay soil textures,while VAR10, VAR20, and VAR30 indicate water-content variability around the nominal value of ±10%, ±20%, and ±30%.

A second ongoing line focuses on material classification in construction environments, in particular the discrimination between drywall and brick structures using the Oliwall platform. Preliminary observations suggest that drywall produces more homogeneous radar responses, whereas brick walls exhibit stronger internal variability, as illustrated in Fig. 6.

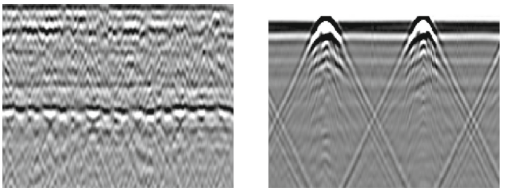


Fig. 6. GPR responses of brick (left) and drywall (right) walls.

5 Experimental Evaluation

The experimental evaluation assesses the proposed processing framework in two representative scenarios. At this stage, data acquisition was performed manually, while following the same sensing configurations and processing pipelines intended for robotic deployment (Fig. 2).

The first experiment addresses root-related signatures in agricultural environments. By comparing excavated root architectures with the corresponding radar responses, large roots are found to produce clear hyperbolic signatures, while additional reflections confirm the non-unique nature of the interpretation. The data was represented as circular slices around the tree, using 10 depth bands between samples 20 and 200 ns and rings with a radial spacing/thickness

of 0.2 m. After standard preprocessing, these slices highlight higher-amplitude zones that correlate with regions of increased root density (Fig. 7), which is often more relevant in practice than exact root reconstruction.

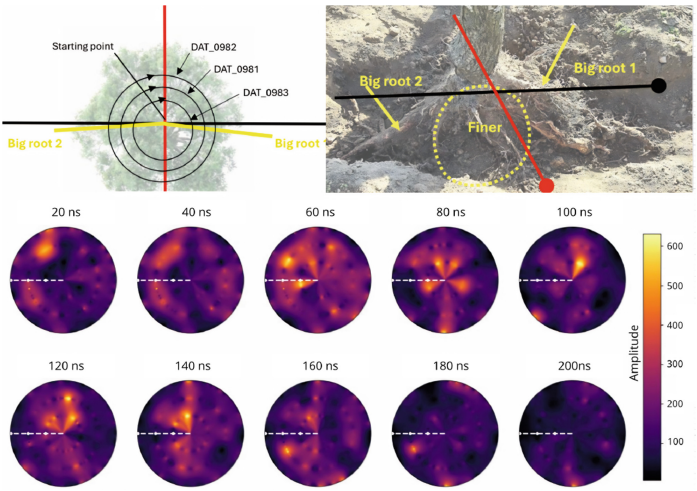


Fig. 7. Comparison between the excavated root layout and circular GPR slices around the tree. The upper panels show the acquisition geometry and excavated roots, while the lower panels show representative amplitude slices. The dashed white line corresponds to the red reference line above. Higher-amplitude regions are spatially consistent with zones of major roots and denser root concentrations. (Color figure online)

The second experiment focuses on material discrimination using buried steel and plastic pipes. Hyperbolas are detected by the AI-based stage, refined by geometric fitting, and analyzed in the frequency domain. In Fig. 8, the radargram spans approximately 60 cm, but the spectral analysis is limited to the first 40 cm: the first two hyperbolas correspond to metallic pipes and the next two to plastic

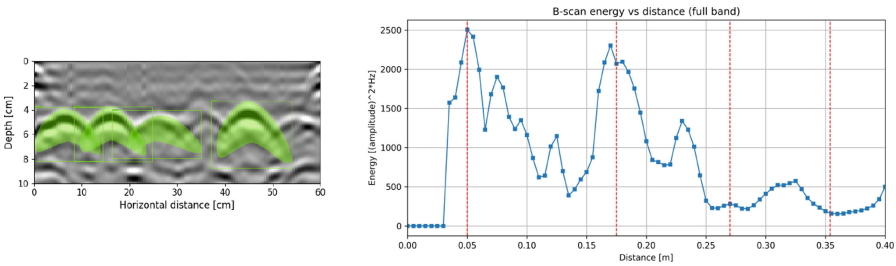


Fig. 8. GPR detections and corresponding energy profile. The radargram extends to 60 cm, while the spectral analysis is limited to the first 40 cm, covering the first four hyperbolas. One plastic target is still poorly detected by the AI model.

pipes, with the last plastic target not detected by the AI model. The remaining detection is not included in the energy profile, as they correspond to another material. The results show that metallic pipes exhibit consistently higher spectral energy than plastic ones, enabling a simple but effective separation.

6 Conclusions and Future Work

The results show that the proposed pipeline is effective for detecting hyperbolic subsurface signatures and supporting application-specific interpretation in different scenarios. In agricultural settings, large roots produce clear hyperbolic responses, while circular slices provide spatial information useful for root density analysis. In controlled construction tests, the combination of AI-based detection, geometric fitting, and frequency-domain analysis enables the estimation of target position and depth, as well as a simple but effective discrimination between metallic and plastic pipes.

More broadly, this work shows that interpretation steps still commonly performed manually can be organized within a robotic GPR framework. Robotic deployment is expected to improve repeatability and coverage, while automated processing reduces the effort required for radargram inspection. At the same time, further automation depends strongly on data availability: some branches already support robust analysis, whereas others still rely on simpler models because larger and more diverse datasets are needed for stronger generalization. This is especially relevant for work-in-progress lines such as soil-moisture estimation and material classification in demolition scenarios.

Future work will focus on full end-to-end robotic deployment. Although the GPR sensors have been integrated with the robotic platforms, the acquisition strategies have not been exploited in real time. The next step is therefore to use the robotic platforms for structured and repeatable scanning, for example in dense root surveys or large wall inspections, and to combine GPR with complementary sensing modalities for more robust and context-aware interpretation.

Acknowledgments. DISCOVER: European Union's Horizon Europe research and innovation programme, grant agreement No. 101129909. ROSAQUA: Government of Catalonia, Agency for the Management of University and Research Grants, Knowledge Industry 2024, grant 2024 PROD 00097. ROOTBOT: Rural Development Program of Catalonia 2014–2022, Department of Climate Action, Food and Rural Agenda, Government of Catalonia.

Disclosure of Interests. The authors have no competing interests.

References

1. Bai, X., et al.: A comprehensive review of conventional and deep learning approaches for ground-penetrating radar detection of raw data. Appl. Sci. **13**(13), 7992 (2023). https://doi.org/10.3390/app13137992

2. Barkataki, N., Tiru, B., Sarma, U.: A CNN model for predicting size of buried objects from GPR B-scans. J. Appl. Geophys. **200**, 104620 (2022). https://doi.org/10.1016/j.jappgeo.2022.104620
3. Bigman, D.P.: Evaluating the feasibility of robotic crawler deployed ground penetrating radar to assess repairs of a concrete hydroelectric dam spillway in Alabama, USA. Water **15**(10), 1858 (2023). https://doi.org/10.3390/w15101858
4. Branco, D., Coutinho, R., Sousa, A., dos Santos, F.: Subsurface metallic object detection using GPR data and YOLOv8-based image segmentation. In: Proceedings of the 21st International Conference on Informatics in Control, Automation and Robotics (ICINCO), pp. 692–699. SciTePress (2024). https://doi.org/10.5220/0013027500003822
5. Caballero, D., et al.: Robotics in sensing the agricultural environment. In: Proceedings of the ASME 2025 International Design Engineering Technical Conferences (IDETC/CIE). Anaheim, CA, USA (2025)
6. Daniels, D.J.: Ground Penetrating Radar. Institution of Engineering and Technology, 2nd edn. (2004)
7. DISCOVER: DISCOVER project, GA 101129909. Horizon Europe Research and Innovation Programme (2024). https://discover-horizon.eu/
8. Eldemiry, A., Muddassir, M., Zayed, T.: Autonomous data acquisition of ground penetrating radar (GPR) using LiDAR-based mobile robot. In: Proceedings of the 41st International Symposium on Automation and Robotics in Construction (ISARC 2024), pp. 206–212. (2024). https://doi.org/10.22260/ISARC2024/0028
9. Fan, G., et al.: 3D reconstruction of root system architecture in urban forest parks based on ground penetrating radar instantaneous amplitude analysis. Plant Phenomics **7**(4), 100147 (2025). https://doi.org/10.1016/j.plaphe.2025.100147
10. GPR Underground Pipeline Detection: GPR scans pipeline dataset. Roboflow Universe (2026). https://universe.roboflow.com/gpr-underground-pipeline-detection/gpr-scans-pipeline. Accessed 18 Apr 2026
11. Halder, S., Afsari, K.: Robots in inspection and monitoring of buildings and infrastructure: a systematic review. Appl. Sci. **13**(4), 2304 (2023). https://doi.org/10.3390/app13042304
12. Huisman, J.A., Hubbard, S.S., Redman, J.D., Annan, A.P.: Measuring soil water content with ground penetrating radar: a review. Vadose Zone J. **2**(4), 476–491 (2003). https://doi.org/10.2136/vzj2003.4760
13. Kalaichelvi, T., Ravi, S., Buddepu, S.: GPR image classification of buried objects using deep learning with attention mechanism. Int. J. Engineering Trends Technol. **73**(4), 24–33 (2025). https://doi.org/10.14445/22315381/IJETT-V73I4P103
14. Lai, W.W.L., Dérobert, X., Annan, A.P.: A review of ground penetrating radar application in civil engineering. NDT & E Int. **96**, 58–78 (2018). https://doi.org/10.1016/j.ndteint.2017.12.006
15. Li, Y., Zhao, Z., Luo, Y., Qiu, Z.: Real-time pattern-recognition of GPR images with YOLO v3 implemented by TensorFlow. Sensors **20**(22), 6476 (2020). https://doi.org/10.3390/s20226476
16. Liso, A., Calderón, N.R., Perez-Gracia, V., D'Orazio, T.R., Renò, V.: Detecting the content of water in agricultural soils by GPR using deep learning. In: Proceedings of SPIE. vol. 13570, p. 3062608 (2024). https://doi.org/10.1117/12.3062608
17. Liu, H., Peng, H., Wang, X., Meng, X., Cui, J., Du, Y.: Comprehensive review of deep learning applications in ground penetrating radar. NDT & E Int. **162**, 103738 (2026). https://doi.org/10.1016/j.ndteint.2026.103738

18. Paul, A., Jayapal, U.M., Ayothiraman, R.: A review of ground penetrating radar for underground utility detection and subsurface profiling: Challenges, strategies and a future-oriented framework. Indian Geotech. J. (2025). https://doi.org/10.1007/s40098-025-01378-1
19. Sagnard, F., Tarel, J.P.: Template-matching based detection of hyperbolas in ground-penetrating radargrams for buried utilities. J. Geophys. Eng. **13**(4), 491–504 (2016). https://doi.org/10.1088/1742-2132/13/4/491
20. Ultralytics: YOLO11 documentation. Online documentation (2024). https://docs.ultralytics.com/models/yolo11/. Accessed 18 Apr 2026
21. Warren, C., Giannopoulos, A., Giannakis, I.: gprMax: open source software to simulate electromagnetic wave propagation for ground penetrating radar. Comput. Phys. Commun. **209**, 163–170 (2016). https://doi.org/10.1016/j.cpc.2016.08.020
22. Zhou, H., et al.: Mapping agricultural soil water content using multi-feature ensemble learning of GPR data. Agricult. Water Manag. **297**, 108884 (2024). https://doi.org/10.1016/j.agwat.2024.108884

Decarbonising Long-Distance Transport: Modal Shift from Short-Haul Aviation to High-Speed Rail for Sustainable Mobility

Oana Dinu[1(✉)], Eugen Rosca[1], Ilona Costea[1], Andrea Alonso Ramos[2], Aura Rusca[1], Anamaria Ilie[1], and Cristian Marin[1]

[1] National University of Science and Technology Politehnica Bucharest, 313 Spl Independentei, District 6, Bucharest, Romania
oana.dinu@upb.ro

[2] Polytechnic University of Madrid, P Juan XXIII, n 11, 28040 Madrid, Spain

Abstract. This paper investigates the sustainable evolution of long-distance transport by examining how rail-oriented strategies can help air travel decarbonization. Focusing on domestic relations in the 300–700 km range, confirmed in the state-of-the-art literature, the study evaluates the potential to shift demand from short and medium-haul aviation to high-speed rail at Romania-wide network scale. The Romanian air and rail transport networks are assessed, quantifications and estimates of traffic are made for airport grid. It analyses connectivity and modal shift potential of high-speed rail line planned from Constanța to the Hungarian border, part of the TEN-T core net, using modal shift scenarios, and estimates CO_2 savings on key routes under optimal rail infrastructure.

Keywords: modal shift · air · high speed rail · connectivity index · decarbonization

1 Introduction

Transport in general, and particularly long-distance transport, are major contributors to polluting emissions. While air transport offers clear advantages, high speed and connectivity, it remains one of the primary sources of pollution. To achieve decarbonization targets across Europe and beyond, air transport faces intense scrutiny, prompting the search for viable alternatives.

Theoretical and practical studies indicate that, for transport distances in the 300–700 km range, short - and medium-haul flights could be partially replaced by rail, especially along shared corridors served efficiently by trains. Statistics points that, on 2024 level, the share of short -haul flights in the distance class of 200–500 km are 2,5% from total passengers, with corresponding CO2 emissions of 1,8% from total emissions and the share of medium -haul flights in the distance class of 500–900 km are 17% from total passengers, with corresponding CO2 emissions of 13% from total emissions [1, 2].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 405–415, 2027.
https://doi.org/10.1007/978-3-032-30539-8_32

The present work builds on this literature by extending the rail–air comparison to a system-level, Romania-wide network. This study adopts a comprehensive methodology combining network analysis of air and rail infrastructure with quantitative connectivity metrics. It employs empirical traffic data to simulate scenarios under varying infrastructure and policy assumptions, quantifying environmental benefits like CO_2 reductions.

In conclusion, the analysis reveals that the potential for modal shift from short-haul aviation to high-speed or conventional rail in Romania remains limited at present. Although plans exist for developing a high-speed rail line spanning approximately 780 km, the current state of infrastructure coupled with the ongoing decline in rail traffic (−50% passenger train-km from 2016–2025), which does not justify major investments does not support a sustainable shift in the foreseeable future.

This paper is structured as follows: in Sect. 2 related work on modal shift and decarbonization is reviewed. Section 3 assesses Romania's air and rail networks. Section 4 analyzes ACI metrics and simulates shift scenarios with CO_2 estimates. Last, in Sect. 5 conclusions, with limitations and future directions, are drawn.

2 Related Work

Decarbonising short-haul aviation is critical for reducing the environmental footprint of the aviation sector, which is a significant contributor to global greenhouse gas emissions. According to Allianz Research [1], each year aviation produces one gigaton of carbon dioxide, which is equivalent to 2.5% of total anthropic emissions. The focus of European Union policies is short-haul flights, which are more frequent and contribute substantially to air pollution and climate change. Sustainable aviation fuels (SAFs), hybrid-electric configurations, and improved operational efficiencies are considered possible solutions for mitigating the carbon footprint of the aviation sector [3–5].

However, in recent years, a modal shift from air to rail transport has gained significant attention due to its potential to reduce environmental impacts and improve transportation efficiency [6, 7]. This modal shift is influenced by environmental, economic and political factors. For certain routes, high-speed rail services have become a competitive alternative to air transport [2].

The modal shift from air to rail transport has benefits and challenges, according to some case studies. Rail transport is considered a sustainable option when powered by clean energy sources [7, 8]. The choice between air and rail transport is determined by factors such as distance and comfort, according to studies using multinomial logit models [9, 10]. Investments in high-speed transport infrastructure improve the development of regions by increasing connectivity, according to studies based on economic models [11]. Integrative frameworks for platform-based models promote climate neutrality [12]. Their principles apply to passenger modal shift, accelerating decarbonization through HSR use. Solutions to improve transit through intermodal passenger terminals promote efficient modal transfers [13]. They strengthen the shift to HSR, optimizing flows at airports like OTP.

Optimizing transport connectivity and efficiency involves network allocation and design. Passenger and freight mobility is improved by integrating transport modes,

thereby reducing congestion, capacity constraints and environmental impacts. Strategies for integrating multimodal transport at different scales and regions, however, have a gap [14]. Some European research projects, such as MODUS [15], have developed macro scenarios for 2040, demonstrating that a short-haul modal shift implies a seamless travel experience. For freight, the transition to rail for high-value goods involves overcoming logistical bottlenecks at intermodal terminals and expediting deliveries [16, 17]. The results of the MODUS project show that an incomplete integration of the fare offer and rail connections at airport terminals prevents the replacement of connecting flights with rail. Good structuring of intermodal networks reduces costs and increases reliability according to some research [18, 19]. Studies focused on Romania are extremely few, making it necessary to study the modernization of conventional railways as an alternative to the missing high-speed transport infrastructure [20]. applies transport forecasting, a key stage in transport management, validating modal shift estimates, projecting realistic HSR demand on regional routes.

The transition from air to rail transport requires the development of policies that facilitate this process. The potential of air-rail intermodality for reducing congestion and decarbonization is evident at transport mode junctions, enabling seamless integration in multimodal hubs [21]. The modal shift to HSR strengthens these interfaces, cutting emissions and boosting accessibility. Increase in the market share of rail transport is encouraged through subsidies and the integration of rail networks with airports [22, 23]. According to the Transport & Environment report [24], this modal shift would reduce intra-European aviation emissions by 4–seven million tons of CO2 per year, a projection modelled on the basis of 3 development scenarios: modernization and increase in speed on rail corridors, 10% reduction in average train travel times, 50% increase in the supply of night trains to carry passengers on long distances. Furthermore, the same report supports the adoption of fiscal policies such as additional taxation on kerosene to encourage the choice of rail transport on distances between 300 and 1000 km. The modal shift from short-haul flights to rail reduces not only CO_2 but also local pollutants like PM2.5. This is especially true in adverse urban atmospheric conditions that amplify traffic-related concentrations. Plus, electrified rail minimizes these non-CO_2 impacts near airports and stations [25]. The shift to HSR boosts positive health effects by promoting clean mobility. Health impact assessments for active travel models in large urban areas show benefits from limiting pollutant emissions [26]

3 Study Area. Assessment of Air and Rail Transport Network

Romania's airport infrastructure constitutes a network of 17 operational facilities, presented in fig. 1, positioned to serve both domestic connectivity and international integration. The system is dominated by Bucharest Henri Coandă International Airport (OTP), which handles over 60% of national passenger traffic, functioning as the primary gateway. Data for the top 10 airports, in total traffic terms, is presented in Table 1.

Regional airports Cluj-Napoca (CLJ), Timișoara (TSR), and Iași (IAS) serve as secondary hubs, while smaller airports serve the niche tourism, cargo, and regional mobility needs.

Romanian airports have undergone substantial modernization, but capacity constraints, despite infrastructure gains, still exist. Service quality in passenger terminals is

key for hub efficiency [27]. The inland connectivity lags, and rail-air intramodality is below EU averages.

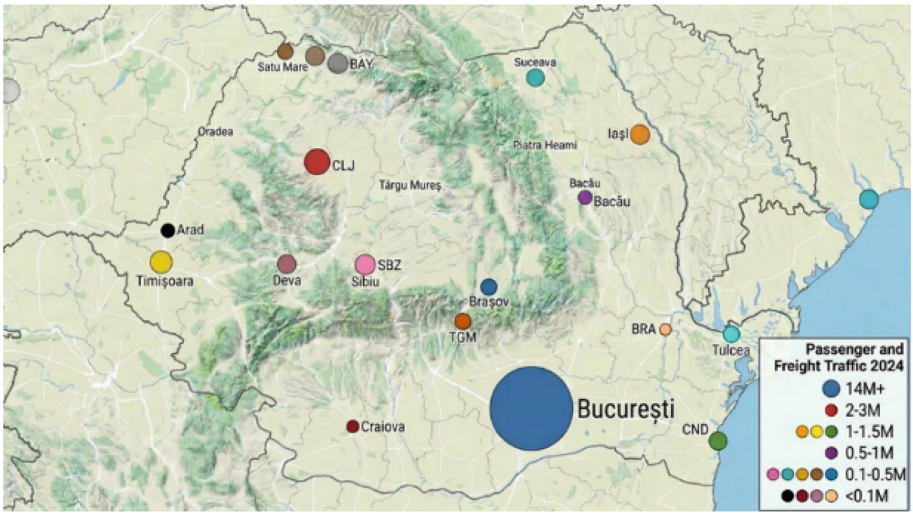


Fig. 1. Romanian top airports in terms of 2024 Passenger and Freight Traffic (*Source: Author's concept)*)

Table 1. Romanian Airports data (2024)

Code	City	Full Name	Traffic [million pax]
OTP	Bucharest	Henri Coandă International Airport	15.9
CLJ	Cluj	Avram Iancu International Airport	2.8
TSR	Timișoara	Traian Vuia International Airport	2.2
IAS	Iași	Iași International Airport	1.5
SBZ	Sibiu	Sibiu International Airport	0.4
BCM	Bacău	George Enescu International Airport	0.6
CRA	Craiova	Craiova International Airport	0.5
OMR	Oradea	Oradea International Airport	0.7
CND	Constanța	Mihail Kogălniceanu Airport	0.8
SUJ	Satu Mare	Satu Mare International Airport	0.2

Romania's rail infrastructure. Analysis of the synthetic characteristics of the Romanian railway infrastructure highlights several limitations that are interesting for the analysis. The share of network length equipped with double track is relatively low, under 30% of

the total network length, representing a limiting factor for the commercial speed of trains. Moreover, the share of network length equipped with electrified track is relatively low, under 40% of the total network length, posing a limiting factor from a decarbonization perspective.

Figure 2 shows rail traffic intensity (trains/day) on Romania's primary network in 2025, with total, passenger, and freight values [28]. For 2025, granulated data per route is unavailable, data from the National Institute of Statistics - INSSE provides national aggregates, Romanian Railways - CFR reports average daily ~1100 passenger trains and, in purple line, the proposed high-speed rail corridor (HSR).

Options for expanding long-distance passenger rail infrastructure fit into strategic sustainable development measures for land transport networks [29]. They validate HSR as a key solution for decarbonizing regional mobility. Romania is preparing a 781.9 km HSR corridor from Constanța to the Hungarian border (see Fig. 2), the opportunity study was completed in December 2025. The alignment runs Constanța – Bucharest – Brașov – Sighișoara (Vânători) – Târgu Mureș – Cluj-Napoca – Zalău – Oradea – border with Hungary, linking to the planned Hungarian high-speed network and forming part of the TEN-T core network.

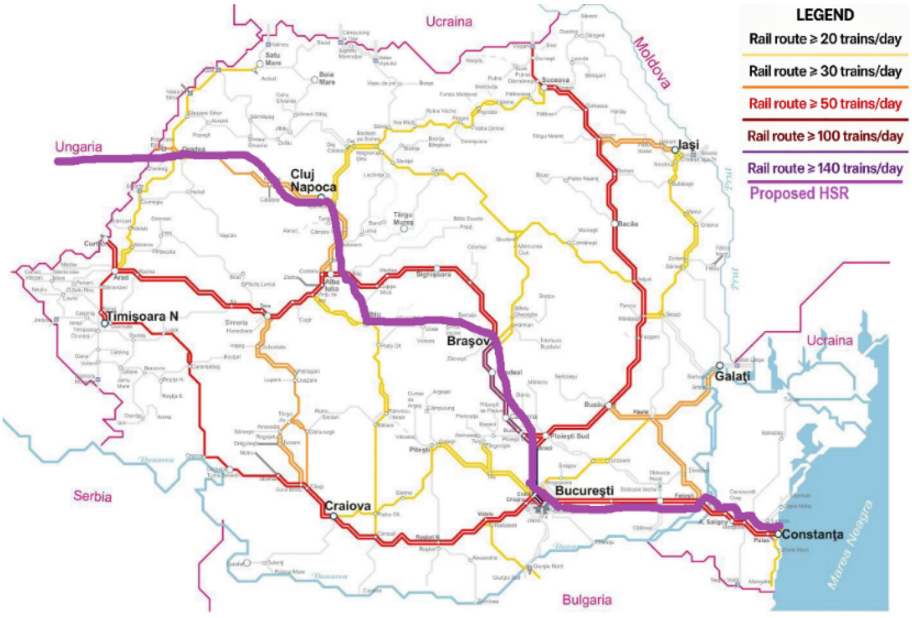


Fig. 2. Romania's existing primary rail network with traffic intensity and proposed HSR (*Source: Author's concept))*

4 Analysis of Potential Air to High-Speed Rail Modal Shift

4.1 Evaluating Air Connectivity

Air connectivity at national level reflects a mix of factors, including economic performance, trade links and openness, diasporas, inbound and outbound leisure demand, and of course geographical position. But air connectivity is also increasingly shaped by national and regional policies, regulation, taxation – and the overall support (or lack of it) for the aviation sector. Romania has now achieved the 2019 level of air connectivity, only −1% for 2025 compared to 2019 level, ranking 12 out of 31 EU+ countries.

In this study, air connectivity is used as a measure of network concentration enabling passenger movement from origin to destination. The direct connectivity denotes the level of connections offered from the assessed airport and indirect connectivity, the level of reasonable connections offered from the assessed airport indirectly through other airports.

From the perspective of measurement metrics, Equation 1, there can be identified key metrics like passenger volumes, fares, direct/indirect destinations, travel times, and hub linkages, all of them network scope and efficiency quantifications.

$$ACI_j = \frac{1}{10000} \sum_{i=1}^{k} f_{ji} \cdot s_{ji} \cdot w_i \tag{1}$$

where:

ACI_j represents air connectivity index for origin airport j,
f_{ji} - flight frequency from j to destination i,
s_{ji} - average seats per flight on route ji,
w_i - weight of destination i,
k - number of destinations,

and the scalar 10000 normalizes units, levelling the index to integer-like values for readability (common in connectivity metrics, yielding ACI ~1000 for EU airports). The weights *wi* were considered 1 for all destinations i prioritizing frequency/seat volume; the distance/population adjustments were initially tested but excluded for ease in this aggregate analysis.

Direct and indirect year-over-year variations of *ACI* from *t* and *t-1* year, in Equation 2, are calculated separately by comparing absolute values calculated annually then computing the percentage change. The traffic values used for calculation are the ones from June peak week.

$$YOY = (ACI_t - ACI_{t-1})/ACI_{t-1} \times 100\% \tag{2}$$

Existing theoretical frameworks and mechanisms underpin the potential modal shift. First, the time-competitive threshold posits that high ACI on routes under 700 km favors HSR when rail travel time approximates air door-to-door duration. Second, price signals amplified by kerosene taxes and low direct ACI encourage shifts if rail offers equivalent frequency of 20 trains/day or more.

Romanian air connectivity metrics. A merged comparison of Romanian airports connectivity, combining historical 2018/2017 data with estimated 2024/2023 from ACI Europe 2024 (Romania total + 3% YoY, direct −0.9%) and 2025/2024 from ACI Europe 2025 (Romania total + 11% YoY, direct −1%) national aggregates, distributed by airport using historical patterns, is presented in Table 2.

Table 2. Romanian Airports Connectivity for direct and indirect traffic

Code	YOY 2018/2017(%)		YOY 2024/2023 (%)		YOY 2025/2024 (%)	
	Direct	Indirect	Direct	Indirect	Direct	Indirect
OTP	2.5	11.7	−0.9	5.2	4.6	7.3
CLJ	−0.5	33.8	5.0	10.0	8.6	12.1
TSR	−6.7	8.8	−2.0	6.0	3.9	11.0
IAS	−6.9	−0.3	3.0	8.0	11.6	4.3
SBZ	12.3	20.5	8.0	4.0	11.3	1.1
BCM	9.0	0.0	12.0	2.0	15.3	3.8
CRA	0.1	0.0	1.0	1.0	9.7	0.9
CND	−5.2	−2.3	−1.0	3.0	1.6	3.4
OMR	27.2	−59.1	15.0	−5.0	21.3	- 0.3
SUJ	49.4	−66.1	20.0	−10.0	25.6	−11.6

The high indirect YOY ACI values for OTP (7.3%) and CLJ (12.1%), estimated for 2025, indicate a pattern of multi-leg travel via hubs for regional access, which HSR can successfully compete with directly door-to-door services under 3 h, on 300–700 km routes. This connectivity configuration justifies the modal shift, as improved rail options reduce the need for indirect air links, attracting traffic according to air-rail intermodality thresholds established in the literature. Most of the airports analysed show signs of indirect recovery, positive indirect trends, in 2024 and 2025, can be observed for most airports. Airports SUJ, OMR show regional operators' volatility, with high fluctuations and instability in operations, networks, routes, and market shares of carriers operating regional short-haul flights. Indirect ACI growth at OTP also reflects feeder-type congestion in airport terminals. Some authors [30] model passenger flows in airport terminals via discrete simulation. This justifies how the HSR shift eases it by diverting regional passengers, reducing terminal bottlenecks. Moreover, air traffic management automation enhances indirect connectivity efficiency at hubs like OTP [31], boosting modal shift potential via competitive HSR

4.2 Air to Rail Modal Shift Scenarios

Post-HSR completion, the connectivity data evaluated in Table 2 enables air-rail shift potential at key airports. Current rail infrastructure constrains this, but the presented scenarios below assume corridor operationalization. Under proper competition conditions, HSR can capture 10–30% share, as simulated in Table 3 scenarios.

The logical link between rising indirect ACI and modal shift tendency has two components. First, indirect connectivity reveals latent demand for seamless regional trips. HSR unlocks this by matching air's total cost, mainly time and fare, via logit-based elasticities. Second, for the 450 km OTP-CLJ route, indirect ACI growth signals substitutability. Although Romania's YOY ACI in 2025 is +11% compared to 2024, the indirect YOY ACI is high for the hubs OTP and CLJ, an indicator that can justify the shift from air to rail.

The hypotheses under which the analysis is conducted are as follows: the high-speed rail corridor from Constanța to the Hungarian border (purple line in Fig. 2) is finalized, the HSR speed is 250 km/h, a level of 10–30% modal split versus air is assumed. The optimistic 30% upper modal split draws from top EU cases like Paris-Lyon (25–35%) and Madrid-Barcelona (20–40%) [3, 4]. These feature high population density and 300 km/h HSR. The sensitivity analysis was adjusted down conservatively for Romania's lower density and 250 km/h HSR.

In those hypotheses, the necessary conditions for considering modal shift are met. They include:

- comparable travel time (door to door),
- the existence of balanced frequency (more than 4 trains/day on peak hours in some studies, or more than 20 trains/day in other studies),
- competitive rail pricing, and
- intermodal facilities for access to the rail infrastructure, with durations of maximum 30 min.

The values used for evaluating items in Table 3 are based on 2024 Romanian air traffic, logit models elasticities, and HSR times from 2025 feasibility study. The estimates of ±30% sensitivity were supposed to be dependent of low traffic/volatility; the maximum 28kt CO_2 saving are conservative vs. EU benchmarks. Emissions of 0.2 kg CO_2/pax-km for air and 0.02 kg CO_2/pax-km for electrified rail were considered for evaluation of decarbonization, data confirmed by ICAO 2025 Environmental Report. Commonly, levels of ~0.15–0.25 kg CO_2/pax-km for short-haul air and 0.01–0.04 kg CO_2/pax-km for electrified rail are used for computations.

Table 3. CO2 saving from HSR modal shift on key relations

Key Relation	Distance (km)	Domestic [M pax/year]	Shifted [M pax/year]	Pax-km [kt]	CO_2 Savings (kt)	Shifted [M pax/year]	Pax-km (kt)	CO_2 Savings (kt)	Shifted [M pax/year]	Pax-km (kt)	CO_2 Savings (kt)
			Scenario 1–10% shift			Scenario 2–20% shift			Scenario 3–30% shift		
OTP-CLJ	450	0.30	0.03	14	2	0.06	27	5	0.09	40	7
OTP-TSR	500	0.20	0.02	10	2	0.04	20	4	0.06	30	5

(*continued*)

Table 3. (*continued*)

Key Relation	Distance (km)	Domestic [M pax/year]	Shifted [M pax/year]	Pax-km [kt]	CO_2 Savings (kt)	Shifted [M pax/year]	Pax-km (kt)	CO_2 Savings (kt)	Shifted [M pax/year]	Pax-km (kt)	CO_2 Savings (kt)
OTP-CND	225	0.10	0.01	2	0.4	0.02	5	1	0.03	7	1
CLJ OTP	450	0.30	0.03	14	3	0.06	27	5	0.09	40	7
TSR OTP	500	0.20	0.02	10	2	0.04	20	4	0.06	30	5
CND OTP	225	0.10	0.01	2	0.4	0.02	4	1	0.03	7	1
Total	–	1.2	0.12	52	**9.8**	0.24	104	**19**	0.36	156	**28**

5 Conclusions

The connectivity analysis performed in Table 2 directly feeds the modal shift scenarios. It uses indirect ACI as a shift potential indicator. Air hubs like OTP/CLJ show growth (7–12% YOY) from feeder traffic.

The estimates are, according to different sensitivity analysis, 0.12–0.36 M shifted passengers/year equivalent to 9.8–28 kt CO_2 savings. CO_2 savings were computed as shifted passenger traffic multiplied with 0.18 kg/pax-km across bidirectional OTP-CLJ, OTP-TSR, and OTP-CND relations. Domestic traffic, (1.2 M pax/year total domestic traffic, aligns with short-haul share) was estimated for each of the key relations as 5–15% of airport totals (e.g., statistics for CLJ report ~9% domestic per 2024 data), focused on OTP links.

OTP-CND traffic is lower than on other key relations due to shorter distance and competition from road/rail, but contributes in some extent to modal shifts

The proposed scenarios, 10% shift, 20% shift and 30% shift apply calculations post-estimation and the results show that savings concentrate on CLJ/OTP (50% of total).

Present study relies on aggregated 2024–2025 traffic data from INSSE/CFR without granular route-level details, for domestic estimates (e.g., 1.2 M pax/year total, 5–15% per route), as considerable data limitations were encountered. The CO2 emission values are consistent with ICAO, but calculations can be refined by considering variability in short-haul operations or rail electrification inconsistencies.

As further development, real life delays and/or costs, validation of scenarios choice for Romanias' case, taking into consideration Romanian road competition, fare integration gaps, and behavioural factors like comfort will allow generalizability beyond selected airports. More empirical surveys and sensitivity tests are needed for robustness as the study innovates Romania-wide network view but underplays economic barriers (e.g., aviation subsidies) and equity risks (e.g., rural access loss).

This contribution informs funding TEN-T policymakers, concerning the much-needed CFR SA upgrades, airport operators (multimodal hubs), and EU (Fit55 targets), advancing sustainable mobility.

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

References

1. Allianz Research: High hopes, heavy footprint: Aviation's quest for climate-neutral skies. In: Economic and Industry Reports (2025)
2. Avogadro, N., Cattaneo, M., Paleari, S., Redondi, R.: Replacing short-medium haul intra-European flights with high-speed rail: impact on CO2 emissions and regional accessibility. Transp. Policy. **114**, 25–39 (2021)
3. ICAO: Environmental Report: Aviation and Environment. International Civil Aviation Organization, Montreal (2025)
4. Jensen, L., Bonnefoy, P., Fitzgerald, J.: The carbon dioxide challenge facing US aviation and paths to achieve net zero emissions by 2050. Prog. Aerosp. Sci. **141**, 100921 (2023). https://doi.org/10.1016/j.paerosci.2023.100921
5. Thonemann, N., Saavedra-Rubio, K., Laurent, A.: Prospective life cycle inventory datasets for conventional and hybrid-electric aircraft technologies. J. Clean. Prod. **434**, 140314 (2024). https://doi.org/10.1016/j.jclepro.2023.140314
6. Baumeister, S., Leung, A.: The emissions reduction potential of replacing short-haul flights with high-speed rail in Europe. Sustainability. **13**(8), 4424 (2021)
7. Prandtstetter, M., Widhalm, P., Starkl, F.: Assessing the potential for a European modal shift towards rail. In: Transport Transitions: Advancing Sustainable and Inclusive Mobility, TRA Conference 2024, vol. 4, pp. 274–284. Springer (2025). https://doi.org/10.1007/978-3-031-95284-5_39
8. Makinen, J., Viri, R., Sjogren, J.: Evaluating passenger transport emissions and modal shift with elasticity-based regional scenario tool. In: Transport Transitions: Advancing Sustainable and Inclusive Mobility, TRA Conference 2024, pp. 181–195. Springer (2025). https://doi.org/10.1007/978-3-031-85578-8_24
9. Chen, Z., Ni, H.W., Wang, Z.L.: Analyzing the propagation effect of high-speed rail from air passenger to air cargo traffic in China. J. Transp. Geogr. **128**, 104335 (2025). https://doi.org/10.1016/j.jtrangeo.2025.104335
10. Nurhidayat, A.Y., Widyastuti, H., Upahita, D.P.: Research on passengers' preferences and impact of high-speed rail on air transport demand. Sustainability. **15**(4), 3060 (2023). https://doi.org/10.3390/su15043060
11. Cao, W.W., Chen, Z.B., Xu, J.: Analysis of travel mode choice behavior between high-speed rail and air transport utilizing large-scale ticketing data. Transp. Res. Rec. **2679**(2), 1344–1358 (2025). https://doi.org/10.1177/03611981241270169
12. Dinu, O. et al.: Integrative framework for platform-based business models to drive climate neutrality in logistics. Sustainability. **16**(4), 1418 (2024). https://doi.org/10.3390/su16041418
13. Roşca, M., Oprea, C., Ilie, A., Olteanu, S., Dinu, O.: Solutions for improving transit through intermodal passenger terminals. Procedia Manuf. **46**, 225–232 (2020)
14. Allard, R.F., Moura, F.: The incorporation of passenger connectivity and intermodal considerations in intercity transport planning. Transp. Rev. **36**(2), 251–277 (2016). https://doi.org/10.1080/01441647.2015.1059379
15. MODUS Project: Future multimodal mobility scenarios within Europe. Deliverable D5.2, Horizon 2020 SESAR Joint Undertaking (2022)
16. Heggen, H., Molenbruch, Y., Braekers, K.: Intermodal container routing: integrating long-haul routing and local drayage decisions. Sustainability. **11**(6), 1634 (2019). https://doi.org/10.3390/su11061634
17. Rezaei, J., Hemmes, A., Tavasszy, L.: Multi-criteria decision-making for complex bundling configurations in surface transportation of air freight. J. Air Transp. Manag. **61**, 95–105 (2017)
18. Fang, Y., Zhang, X., Tao, Y.: Modeling railway express freight train operation plans in road-rail multimodal transport. IEEE Access. **13**, 145996–146007 (2025). https://doi.org/10.1109/ACCESS.2025.3600470

19. Yan, X., Nie, L., Mei, Z.: Optimization design of space connection scheme for air-rail intermodal transport. In: Eighth International Conference on Traffic Engineering and Transportation System (ICTETS 2024), pt, vol. 1, p. 13421. SPIE (2024). https://doi.org/10.1117/12.3054813
20. Dragu, V., Dinu, O., Oprea, C., Roman, E.A.: The transport forecast - an important stage of transport management. In: International Congress of Automotive and Transport Engineering - Mobility Engineering and Environment (CAR2017), vol. 252, p. 12050 (2017). https://doi.org/10.1088/1757-899X/252/1/012050
21. Oprea, C., Dragu, V., Ilie, A., Dinu, O., Burciu, Ş.: Transport modes junction in an airport. In: 11th International Conference Interdisciplinarity in Engineering, INTER-ENG 2017, vol. 22, pp. 878–885 (2018). https://doi.org/10.1016/j.promfg.2018.03.124
22. Behrends, S.: Burden or opportunity for modal shift? Embracing the urban dimension of intermodal road-rail transport. Transp. Policy. **59**, 10–16 (2017). https://doi.org/10.1016/j.tranpol.2017.06.004
23. Xia, W.Y., Wang, K., Zhang, A.M.: Air transport and high-speed rail interactions in China: review on impacts of low-cost carriers, rail speed, and modal integration. Airline Econ. Asia. **7**, 103–122 (2019). https://doi.org/10.1108/S2212-160920180000007007
24. Transport & Environment: Air2Rail: reducing CO2 from Intra-European Aviation by a Modal Shift from Air to Rail. T&E Reports, Brussels (2020)
25. Roșca, E. et al.: A moderation effect of atmospheric conditions on PM2.5 traffic-related concentration. Rom. J. Transp. Infrastruct. **12**(1), 1–17 (2023). https://doi.org/10.2478/rjti-2023-0007
26. Ilie, A., Oprea, C., Ruscă, A., Dinu, O., Olteanu, S., Burciu, Ş.: Health impact assessment study on active travel in large urban areas. In: 12th International Conference Interdisciplinarity in Engineering (INTER-ENG 2018), vol. 32, pp. 944–951 (2019). https://doi.org/10.1016/j.promfg.2019.02.307
27. Oprea, C., Roșca, E., Popa, M., Ilie, A., Dinu, O., Roșca, M.: The quality of service in passenger transport terminals. In: 20th Innovative Manufacturing Engineering and Energy Conference (IMANEE 2016), vol. 161, p. 12098 (2016). https://doi.org/10.1088/1757-899X/161/1/012098
28. Ministry of Transport and Infrastructure, Compania Națională de Căi Ferate "CFR" SA: Annex to the Railway Infrastructure Development Strategy 2021–2025, Bucharest (2020) https://cfr.ro/strategia-de-dezvoltare-a-infrastructurii-feroviare/. Accessed 30 Mar 2026
29. Dinu, O., Burciu, Ş.: Options for long distance passenger land transport infrastructure expansion. Analysis regarding network-based sustainable development. In: 9th International Conference Interdisciplinarity in Engineering, INTER-ENG 2015, vol. 22, pp. 991–998 (2016). https://doi.org/10.1016/j.protcy.2016.01.124
30. Oprea, C. et al.: Analyzing passenger flows in an airport terminal: a discrete simulation model. Computation. **12**(11), 223 (2024). https://doi.org/10.3390/computation12110223
31. Donciu, B., Constantinescu, C.E., Pleter, O.T.: A case study in air traffic management automation. UPB Sci. Bull. Ser. D. **73**(3), 63–70 (2011)

Nanofluid Injection in Poroelastic Media: Possible Applications in Environmental Science

Rosella Sampalmieri[1], Federica Di Michele[2], Donato Pera[1(✉)], and Bruno Rubino[1]

[1] Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila, via Vetoio, loc. Coppito 1, L'Aquila, Italy
donato.pera@univaq.it

[2] Istituto Nazionale di Geofisica e Vulcanologia, via Alfonso Corti 12, Milan, Italy

Abstract. Nanofluids are innovative materials consisting of a suspension of nanoparticles in a fluid, such as water or oil. They are commonly used in engineering and medicine because of their thermal conductivity, which makes them particularly efficient for cooling, heat transport, and tumor treatment. In recent times, the unique properties of nanofluids have led to their use for environmental application. This includes oil extraction processes, the treatment of contaminated water and soil, and the improvement of the seismic response of the soil during strong seismic events. Although these applications are very different from each other, it is necessary in all cases to inject the nanoparticles into the soil. However, predicting the distribution of particles *a priori*, considering the injection parameters, is very difficult due to the intrinsic complexity of the subsoil and uncertainty regarding parameters such as porosity and elastic properties. One possible solution is to use numerical simulations on digital twins of the area of interest. These simulations are computationally expensive, and are affected, in the specific case, by a high level of uncertainty, due to the lack of specific information on soil stratigraphy.

In this paper, we present and discuss a mathematical model that describes the injection of nanofluid into the soil, taking into account its natural poroelasticity, which is often neglected. The subsequent drift and diffusion of the particles, as well as their interaction with the surrounding medium through absorption and desorption processes, can be described using the same model, but with different boundary and initial conditions. After introducing the model, we discuss in detail the effects of poroelasticity and the role of injection parameters on particle distribution. We also present some (*semi*−)analytical representations of the model that could be useful for the preliminar calibration of the numerical simulations and, potentially, of the injection process.

Keywords: Nanofluid · porous media · pollution · fluid injection · poroelasticity · nanoparticles

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 416–426, 2027.
https://doi.org/10.1007/978-3-032-30539-8_33

1 Introduction

Nanofluids are aqueous (or other fluid) suspensions containing small percentages (max. 1% by volume) of appropriately functionalized nanoparticles (NPs) [5]. Nanofluids are widely used in numerous fields of scientific and technological research thanks to their unique chemical and physical properties and relatively low production cost. Because of anthropogenic pollution, the development of nanotechnologies is regarded as a beneficial environmental tool, especially in the context of in situ remediation, that is, the cleanup of soils and groundwater [3,9]. The enhanced penetration capacity of NPs, attributable to their reduced size, facilitates their injection into the matrix at depth, thereby facilitating the degradation and immobilization of contaminants across a wide range of sites. In addition, the implementation of NPs is often more cost-effective and less destructive than that obtained with ex situ techniques. However, the biological risk associated with the use of these techniques has yet to be fully assessed, given their recent application [13]. Anyway, NPs technology also has many other applications in several fields, some of which are medicine (because their size is similar to that of bio-molecules), cosmetics, restoration and preservation of drinking water supplies, cooling electronics, structural engineering, nuclear engineering, oil industry and injection of nanofluids into the soil to enhance soil properties for agricultural purposes [10] or for improving soil stability [4]. We can also consider emerging trends in green synthesized and bio-based nanoparticles with the aim of safe remediation of the environment [9,14]. In this work, during the injection phase, we reasonably assume that our matrix soil behaves as a poroelastic medium introducing the volumetric dilation in our model. The controlled injection of nanofluids into poroelastic matrices has many important applications, such as in biomedical engineering (drug delivery, hyperthermia cancer treatments [2,7]). Moreover, poroelastic models help predict fluid transport, tissue deformation, and the distribution of nanoparticles, which is critical for targeting therapy effectively and minimizing harm to surrounding healthy tissue. One of the leading application is the nanofluid-enhanced oil recovery: an advanced technique ([11]) used to improve oil production from reservoirs by reducing the interfacial tension between oil and water and increasing rock's wettability (to a more water-wet state). In geothermal systems nanofluids can improve the thermal conductivity of the heat transfer fluid. The process of injecting nanofluids into the soil and their subsequent interaction with the surrounding environment is a highly complex process that depends on numerous factors. Some of these, such as injection parameters, can be calibrated and optimized, while others, such as the geological structure of the sub-soil and the actual porosity of the soil, are very difficult to determine with precision. Hence, it can be very useful to develop accurate mathematical and numerical models capable of supporting the planning phase of NPs injection to maximize the beneficial effect and contribute to cost reduction. On the other hand, thanks to the growing availability of computational resources, it is now possible to simulate complex events on digital twins of real geological domains. However, the accuracy and reliability of these simulations are strongly related to knowledge of the geological structure of the region

under examination, which is often only partially available. Analytical or semi-analytical solutions, even if related to simplified regimes and models, provide an easy, straightforward guidance, that can be used to prepare numerical simulations, but can also help with the design stages of the injection process, reducing the costs of the injection process and optimizing its effectiveness. In this paper, we present some preliminary results of our work aimed at determining the concentration of nanoparticles following the injection process into the subsoil. In this study we do not consider a particular geological structure. However, to set a realistic context, we can consider as soil model a quartz sand matrix and as NPs model, silica-based nanoparticles, which are among the most promising materials for the ease with which it is possible to modify their characteristics. The paper is organized as follows. In Sect. 2 we introduce the mathematical model able to describe the injection process and the successive phase. In Sect. 3 we discuss the procedure for equation scaling, needed to understand which terms govern the process. In Sect. 4 some simple analytical solutions are derived and discussed to understand the relative contribution of the different parameters during and after the injection phase. In Sect. 5 some conclusions and future developments are presented and discussed.

2 Mathematical Model

In this section, we briefly introduce the mathematical model based on the colloid filtration theory [1,12]. The single-phase nanoparticles flow in three-dimensional domains can be described by the reaction-convection-diffusion equation, as follows:

$$\begin{cases} \dfrac{\partial C}{\partial t} + \nabla \cdot \left(\mathbf{v} \dfrac{C}{\varepsilon} \right) = \nabla \cdot \left(D\, \varepsilon\, \nabla \left(\dfrac{C}{\varepsilon} \right) \right) - k_f C \\ \nabla \cdot \mathbf{v} = 0 \\ \dfrac{\partial C_b}{\partial t} = k_f C \end{cases} \tag{1}$$

where $\mathbf{x} \in R^3$, and $t \geq 0$. C and C_b are the concentration (the number of nanoparticles moles per unit volume of suspension) of suspended unbounded particles and of the bounded ones (i.e. adsorbed) respectively, and $\mathbf{v}$ is the Darcy velocity. D and ε denote the diffusion coefficient and the soil porosity. In addition to the mechanical flow terms, the reaction term, that accounts for the interactions between the NPs and the porous medium, represents the NPs adsorption and it is derived from the classical colloid filtration theory. It is modeled through the filtration rate coefficient as $k_f = \frac{3\eta}{4r_c} \frac{1-\varepsilon}{\varepsilon} \|\mathbf{v}\|$. Roughly speaking the coefficient k_f models how fast the NPs can be captured by the soil grains. It depends on the velocity of the fluid that transports the particles, the porosity, the grain size r_c and the efficiency of the collector η. We observe that since the internalization and desorption phenomena, occur over a longer time-scale, compared to the injection phase, their effect can be neglected at this stage. The equations above must be coupled to the Darcy's law modeling the fluid flow in a porous medium: $\nabla p = -\frac{\mu}{K}\mathbf{v}$. Here p is the pressure, μ the

dynamical viscosity and K the soil permeability. Classical models that deal with transport of NPs in the soil describe the medium as porous. However, especially during the injection phase, it is reasonable to assume that the soil behaves as a poroelastic medium. We can therefore assume that porosity ε varies with respect to the rest value ε_0, according to the following law [6]

$$\varepsilon = \frac{\varepsilon_0 + e}{1 + e}, \tag{2}$$

where $e = \nabla \cdot u$ is the volumetric dilation modeled as the divergence of the soil displacement $\mathbf{u}$. We consider $\mathbf{u}$ governed, in a linear regime, by Navier's equation as follows

$$G\,\nabla^2 \mathbf{u} + (\lambda + G)(\nabla e) = \nabla p. \tag{3}$$

Here G and λ are the Lamé parameters. In order to calculate the analytical or semi−analytical solutions, we need to simplify the model. Therefore, we consider the soil as an isotropic and homogeneous medium. Obviously, this is a quite strong approximation, under which the system (1) and the relative closure relations can be written in spherical coordinates, assuming radial symmetry. As explained in the introduction, also the analytical solution of this simple model can provide an estimate, albeit approximate, of the distribution of particles during injection and contribute to the careful selection of injection parameters, which are necessary for the validation of the numerical models. First of all, we observe that for spherically symmetric deformation, (3) simply reads:

$$\alpha \nabla\, e = \nabla\, p, \tag{4}$$

where $\alpha = (\lambda + 2G)$. Then (4) can be integrated with respect to r in the interval $[0, +\infty)$, obtaining

$$e = \frac{p - p_0}{\alpha}. \tag{5}$$

Here p_0 is the pressure in the matrix soil and $e = 0$ before the injection. In the assumption of spherical symmetry, Darcy's law reads as $\frac{\partial p}{\partial r} = -\frac{\mu}{K} v$, where gravity has been neglected in this approximation. The soil permeability K, [8], can be expressed as $K = K_0 \exp(\gamma e)$, where γ is the constant that represents the sensitivity of permeability to volumetric strain e and K_0 is the permeability of the soil in the equilibrium condition. Then, in view of the previous model, the variables to consider are: $C(r,t)$, $v(r,t)$, $p(r,t)$, and $C_b(r,t)$. In the new coordinate system, the equations read as

$$\begin{cases} \dfrac{\partial C}{\partial t} + v \dfrac{\partial}{\partial r}\left(\dfrac{C}{\varepsilon}\right) = \dfrac{1}{r^2}\dfrac{\partial}{\partial r}\left(r^2 D \varepsilon \dfrac{\partial}{\partial r}\left(\dfrac{C}{\varepsilon}\right)\right) - k_f C, \\ \dfrac{\partial}{\partial r}(r^2 v) = 0, \\ \dfrac{\partial p}{\partial r} = -\dfrac{d_0}{K_0 \exp(\gamma e)} v, \\ \dfrac{\partial C_b}{\partial t} = k_f v C, \end{cases} \tag{6}$$

where $r \in R^+$, is the radius of our spherical domain S, d_0 is the NPs diameter and v now denotes the radial component of the Darcy velocity $\mathbf{v}$, which is positive in our setting.

3 Scaling of the Equation and Spherical Symmetrically Analytical Solution

Before calculating the analytical solution, it is necessary to scale the equations, in order to make any unknown dimensionless. In this section we shortly discuss the procedure for equation scaling, needed to understand which terms govern the process. This is a central step in our analysis, as it allows us to solve the problem analytically, simplifying the possible, subsequent, numerical analysis.

During the injection phase, the nanofluid is pumped into the soil, and, in a first approximation, diffusion term can be neglected (advection-dominated regime), because the movement of particles is mainly due to advection. However, at this stage, because of the high-pressure gradients induced by the injection of the nanofluid, the soil can undergo non-negligible deformations and it behaves as a poro-elastic medium. The variation of porosity is assigned according to (2) and (3). After the injection phase, the NPs diffuse into the soil and interact with the soil grains. In reality, the interaction process is very complex and difficult to model due to the heterogeneity of the substrate in which the particles move and the variability of the conditions caused, for example, by the presence of aquifers. In our paper two different regimes are analyzed and discussed separately: the injection phase and the subsequent phase, when the injection stops. We just point out that, in the present analysis, a homogeneous matrix of quartz sand was chosen as porous medium model, in order to fix realistic parameters. This choice does not represent the composition of a real soil and, for sure, it will influence the dynamics of the problem. In a later work we will discuss a more general model where different materials and a non-homogeneous soil will be considered in a more realistic 3D digital twin of the computational model. To obtain the scaled governing equations we need to introduce the relative reference quantities. In more detail we consider t_s as the characteristic injection duration, v_s as the characteristic fluid speed, C_s is the reference NPs concentrations and p_s is the reference pressure. Then we introduce the following scaled variables, normalizing each quantity with respect to its reference values:

$$\bar{t} = \frac{t}{t_s}, \quad \bar{r} = \frac{r}{t_s v_s}, \quad \bar{v} = \frac{v}{v_s}, \quad \bar{p} = \frac{p - p_0}{p_s}, \quad \bar{C} = \frac{C}{C_s}, \quad \bar{C}_b = \frac{C_b}{C_s}.$$

The value C_s is the reference NPs concentration corresponding to a reference value of the pressure p_s. The bar above the letters indicates the scaled variables.

In the scaled variables system (6) reads

$$\begin{cases} \dfrac{\partial \bar{C}}{\partial \bar{t}} + \bar{v}\dfrac{\partial}{\partial \bar{r}}\left(\dfrac{\bar{C}}{\varepsilon}\right) = \dfrac{1}{\bar{r}^2}\dfrac{\partial}{\partial \bar{r}}\left(\bar{r}^2 D^* \varepsilon \dfrac{\partial}{\partial \bar{r}}\left(\dfrac{\bar{C}}{\varepsilon}\right)\right) - k_f^* \bar{C}, \\ \dfrac{\partial}{\partial \bar{r}}(\bar{r}^2 \bar{v}) = 0, \\ \dfrac{\partial \bar{p}}{\partial \bar{r}} = -\dfrac{d^*}{\exp(\gamma e)}\bar{v}, \\ \dfrac{\partial \bar{C}_b}{\partial \bar{t}} = k_f^* \bar{v} \bar{C}, \end{cases} \tag{7}$$

where $\bar{r} \in R^+$, $D^* = \frac{D}{t_s v_s^2}$, $k_f^* = \frac{(3\,\eta\, t_s\, v_s)(1-\varepsilon)}{4\,\varepsilon\, r_c}$, $d^* = \frac{\eta\, t_s v_s^2}{p_s K_0}$, $\varepsilon = \frac{\varepsilon_0 + \frac{p-p_0}{\alpha}}{1+\frac{p-p_0}{\alpha}}$. Note that from the scaling we can write

$$e = \frac{\bar{p}\, p_s}{\alpha} \quad \text{andthen, from (2), wehave} \quad \varepsilon = \frac{\varepsilon_0 + \frac{\bar{p} p_s}{\alpha}}{1 + \frac{\bar{p} p_s}{\alpha}}. \tag{8}$$

In the previous formulas r_c denotes the typical soil grain radius and the collector efficiency η, which typically refers to the fraction of nanoparticles that are effectively captured or deposited on the sand grain surfaces during the flow, is assumed constant. C_b, the bounded NPs concentration, is assumed to be negligible at the beginning of the injection process. We just point out that we have already exploited the incompressibility of the Darcy velocity field.

4 Injection Phase

In this subsection our aim is to model the injection phase, in which the diffusion phenomenon plays a limited role, therefore we will remove the diffusion term from system (7) and we obtain

$$\begin{cases} \dfrac{\partial \bar{C}}{\partial \bar{t}} + \bar{v}\dfrac{\partial}{\partial \bar{r}}\left(\dfrac{\bar{C}}{\varepsilon}\right) = -k_f^* \bar{C}, \\ \dfrac{\partial}{\partial \bar{r}}(\bar{r}^2 \bar{v}) = 0, \\ \dfrac{\partial \bar{p}}{\partial \bar{r}} = -\dfrac{d^*}{\exp(\gamma e)}\bar{v}, \\ \dfrac{\partial \bar{C}_b}{\partial \bar{t}} = k_f^* \bar{v} \bar{C}. \end{cases} \tag{9}$$

To this system we need to add the following suitable initial boundary conditions

$$\begin{cases} \bar{C}(\bar{r}, 0) = C_0(\bar{r}), \\ \bar{C}(r_i, t) = C_i \\ \bar{p}(r_i, t) = p_i \\ 4\pi\, r_i^2\, v(r_i, t) = Q_i \\ \bar{C}_b(\bar{r}, 0) = 0, \end{cases} \tag{10}$$

where $\bar{r} \in R^+$, Q_i is the infusion flow rate, C_i is the initial nanofluid concentration in the injection point, $r_i = r(0)$, is the radius of the injection area, p_i is the pressure in the injection point and p_0 is the value of the soil pressure before the injection ($p_i \geq p_0$). Here p_i, C_i, and r_i can be considered as design parameters. Within the constraints of current technological limits, these variables can be tuned to optimize the injection process and enhance the overall delivery efficiency. We note also that we assume $C_0(\bar{r})$ small but different from zero since we mean to apply the characteristics method to try to find an analytical solution to $(9)_1 - (10)_{1,2}$. Indeed, as a first order hyperbolic PDE, the transport equation $(9)_1$ can be solved with this method which, instead of looking at the entire surface $C(r,t)$, reconstruct the solution along the trajectories of the single particle of the nanofluid, called characteristic curves. Taking the total derivative of C along the characteristics the partial differential equation becomes an ordinary differential equation. In the sequel, to simplify the notation, we will drop the bar above the scaled variables. Note that from (2) and (5) we have that $\varepsilon = \varepsilon(e(p))$ and $K = K(e(p))$. We start from $(9)_2$and $(10)_4$ and we get v is a function of r only as $v(r) = \frac{Q_i}{4\pi\, r^2}$. While, recalling (8), we can write $(9)_3$, as $\frac{\partial p}{\partial r} = -\frac{d^* Q_i}{4\pi\, r^2} \exp\left(-\frac{\gamma p_s}{\alpha p}\right)$. By separation of variables, taking into account the boundary condition $p(r_i, t) = p_i$, we obtain the explicit solution

$$p(r) = \frac{\alpha}{\gamma p_s} \ln\left[\exp\left(\frac{\gamma p_s p_i}{\alpha}\right) + \frac{\gamma p_s d^* Q_i}{4\pi\alpha}\left(\frac{1}{r} - \frac{1}{r_i}\right)\right]. \tag{11}$$

We point out that the pressure is also independent of time. Concerning the transport equation for the nanofluid concentration, $(9)_1$ with $(10)_1$, we will distinguish three different situations depending on the generality with which the porosity ε is chosen, as discussed in the following subsections.

4.1 General Porosity: $\varepsilon(r)$

According to the expression of the quantities (2), (5) and (11), we have that the general formulation for the porosity is just a function of r. Therefore, equation $(9)_1$ can be written as $\frac{\partial C(r,t)}{\partial t} + \frac{Q_i}{4\pi\, r^2}\frac{\partial}{\partial r}\left(\frac{C(r,t)}{\varepsilon(r)}\right) = -k_f^*(\varepsilon(r))C(r,t)$ and then $\frac{\partial C(r,t)}{\partial t} + \frac{Q_i}{4\pi\, r^2\, \varepsilon(r)}\frac{\partial C(r,t)}{\partial r} = \left[-k_f^*(\varepsilon(r) + \frac{Q_i}{4\pi\, r^2}\frac{\varepsilon'(r)}{\varepsilon^2(r)}\right] C(r,t)$. Applying the characteristic method we have that characteristic curves equations are

$$(\dot{r}(s) = \frac{Q_i}{4\pi r^2 \varepsilon(r)}, \dot{t}(s) = 1) \quad \text{from which we have } \int_{r(0)}^{r(t)} 4\pi\, \rho^2\, \varepsilon(\rho) d\rho = Q_i\, t.$$

Let $V(x)$ be a differentiable, strictly increasing function such that $V'(x) = 4\pi\, x^2\, \varepsilon(x)$, then from $V(r(t)) - V(r(0)) = Q_i\, t$ it follows $r(t) = V^{-1}\left[V(r_i) + Q_i t\right]$ where $r_i = r(0)$. Along the characteristic curves $r(t)$, the material derivative of the concentration reveals an additional term proportional to the porosity gradient $\varepsilon'(r)$. This term accounts for the local compression or expansion of

the fluid volume as it moves through a medium of varying porosity, effectively modulating the filtration rate k_f^* as $\frac{dC(r(t),t)}{dt} = \left[-k_f^*(\varepsilon(r) + \frac{Q_i}{4\pi\, r^2}\frac{\varepsilon'(r)}{\varepsilon^2(r)} \right] C(r,t)$,

from which

$$C(r,t) = C_0(r(0)) \exp\Big[\int_0^t \left(-k_f^*(\varepsilon(r)) + \frac{Q_i}{4\pi\, r^2}\frac{\varepsilon'(r)}{\varepsilon^2(r)} \right) ds \Big], \tag{12}$$

that is

$$C(r,t) = C_0(V^{-1}\big[V(r(t)) - Q_i t\big]) \exp\Big[\int_0^t \left(-k_f^*(\varepsilon(r) + \frac{Q_i}{4\pi\, r^2}\frac{\varepsilon'(r)}{\varepsilon^2(r)} \right) ds \Big].$$

From $(9)_4$ and $(10)_5$ we get $C_b(r,t) = \frac{Q_i}{4\pi\, r^2} \int_0^t k_f^*(\varepsilon(r,s))\, C(r,s) ds$.

4.2 Linearized Porosity

In this subsection, with the aim of finding a closed form for the concentration of the nanofluid, let's assume that the value of the volumetric dilation e is small. This hypothesis allows to approximate (2) as

$$\varepsilon \approx \varepsilon_0 + (1-\varepsilon_0)e = \varepsilon_0 + \frac{1-\varepsilon_0}{\alpha}(\bar{p} - p_0) = \varepsilon_0 + \frac{1-\varepsilon_0}{\alpha}(p\, p_s).$$

Moreover, we simplify also the pressure equation in $(9)_3$ assuming $\gamma \approx 0$ into the permeability expression. Then, instead of $(9)_3$, we have $\dfrac{\partial p}{\partial r} = -\dfrac{d^* Q_i}{4\pi r^2}$ with $p(r_i,t) = p_i$ and the corresponding solution is $p(r) = p_i + \frac{d^* Q_i}{4\pi}\left(\frac{1}{r} - \frac{1}{r_i}\right)$.

To solve the equation for concentration

$$\frac{\partial C(r,t)}{\partial t} + \frac{Q_i}{4\pi\, r^2}\frac{\partial}{\partial r}\Big(\frac{C(r,t)}{\varepsilon(r)}\Big) = -k_f^*(\varepsilon(r)) C(r,t), \tag{13}$$

we set $U(r,t) = \frac{C(r,t)}{\varepsilon(r)}$, obtaining $\varepsilon(r)\frac{\partial U}{\partial t} + \frac{Q_i}{4\pi r^2}\frac{\partial U}{\partial r} = -k_f^* \varepsilon(r) U$. Then, we divide by $\varepsilon(r)$ and apply the characteristic method to

$$\frac{\partial U}{\partial t} + \frac{Q_i}{4\pi r^2 \varepsilon(r)}\frac{\partial U}{\partial r} = -k_f^* U.$$

Working similarly as in the previous subsection we have as solution

$$C(r,t) = \frac{\varepsilon(r)}{\varepsilon(r_i)} C_0 \left(t - \int_{r_i}^r \frac{4\pi\rho^2 \varepsilon(\rho)}{Q_i} d\rho \right) \exp\left(-\frac{3\eta t_s v_s}{4\, r_c} \int_{r_i}^r \frac{(1-\varepsilon(\rho))}{\varepsilon(\rho))} d\rho \right)$$

but this time we can compute explicitly the integrals that appear. Indeed, substituting the simplified value of the pressure $p(r) = p_i + \frac{d^* Q_i}{4\pi}\left(\frac{1}{r} - \frac{1}{r_i}\right)$ in

$\varepsilon(r) = \varepsilon_0 + \frac{1-\varepsilon_0}{\alpha}(p(r)\,p_s)$, we have $\varepsilon(r) = \varepsilon_0 + \frac{(1-\varepsilon_0)p_s}{\alpha}\left(p_i + \frac{d^*Q_i}{4\pi}\left(\frac{1}{r} - \frac{1}{r_i}\right)\right)$. We first consider

$$\int_{r_i}^{r} \frac{4\pi\rho^2\varepsilon(\rho)}{Q_i}d\rho = \frac{4\pi}{Q_i}\int_{r_i}^{r} \rho^2\left[\varepsilon_0 + \frac{(1-\varepsilon_0)p_s}{\alpha}\left(p_i + \frac{d^*Q_i}{4\pi}\left(\frac{1}{\rho} - \frac{1}{r_i}\right)\right)\right]d\rho.$$

The result is a combination of powers of order three and two of r and r_i. Concerning the second integral, we have

$$\int_{r_i}^{r} \frac{1-\varepsilon(\rho)}{\varepsilon(\rho)}d\rho = \left(\frac{1-A}{A}\right)(r - r_i) - \frac{B}{A^2}\ln\left|\frac{Ar+B}{Ar_i+B}\right|$$

where $A = \varepsilon_0 + \frac{(1-\varepsilon_0)p_s}{\alpha}\left(p_i - \frac{d^*Q_i}{4\pi r_i}\right)$ and $B = \frac{(1-\varepsilon_0)p_s d^*Q_i}{4\pi\alpha}$. In this framework, the variable porosity $\varepsilon(r)$ acts as a structural constraint that dictates both the local front velocity and the particle capture efficiency, thereby coupling nanofluid transport directly to the mechanical deformation of the medium.

4.3 Constant Porosity

In this subsection we consider the behavior of the concentration of the nanofluid when the volumetric dilation can be assumed negligible, then we can set a constant porosity that equals the original one before the injection, that is ε_0. Then equation $(9)_1$ becomes $\frac{\partial C}{\partial t} + \frac{v}{\varepsilon_0}\frac{\partial C}{\partial r} = -k_f^* C$ and recalling $v(r) = \frac{Q_i}{4\pi r^2}$ and $k_f^* = \dfrac{(3\,\eta\,t_s\,v_s)(1-\varepsilon_0)}{4\,\varepsilon_0\,r_c}$, we have

$$\frac{\partial C}{\partial t} + \frac{Q_i}{4\pi\,r^2\,\varepsilon_0}\frac{\partial C}{\partial r} = -\frac{(3\,\eta\,t_s\,v_s)(1-\varepsilon_0)}{4\,\varepsilon_0\,r_c}\,C. \tag{14}$$

Applying the characteristic method we have that the characteristic curves equations are

$$\begin{cases} \dot{r}(s) = \frac{Q_i}{4\pi\,r^2\,\varepsilon_0} \\ \dot{t}(s) = 1, \end{cases}$$

from which we have $r(t) = \sqrt[3]{r_i^3 + \frac{3Q_i\,t}{4\pi\,\varepsilon_0}}$ and so the trajectory that passes through $(r(t), t)$ is related to the initial position by the relation $r(0) = r_i = \sqrt[3]{r(t)^3 - \frac{3Q_i\,t}{4\pi\,\varepsilon_0}}$. From (14) one has $\frac{dC}{dt} = -\dfrac{(3\,\eta\,t_s\,v_s)(1-\varepsilon_0)}{4\,\varepsilon_0\,r_c}\,C$ and then

$$C(r,t) = C_0(r_i)\exp\left(-\int_0^t \frac{(3\,\eta\,t_s\,v_s)(1-\varepsilon_0)}{4\,\varepsilon_0\,r_c}\,ds\right),$$

that is $C(r,t) = C_0\left(\sqrt[3]{r(t)^3 - \frac{3Q_i\,t}{4\pi\,\varepsilon_0}}\right)\exp\left(-\dfrac{(3\,\eta\,t_s\,v_s)(1-\varepsilon_0)}{4\,\varepsilon_0\,r_c}\,t\right)$. Clearly, once $C(r,t)$ is determined, the concentration $C_b(r,t)$ of the bounded nanoparticles can be easily obtained from $C_b(r,t) = \frac{Q_i}{4\pi\,r^2}\int_0^t k_f^*(\varepsilon(r,s))\,C(r,s)ds$. In this

simplified scenario, the expression for $C(r,t)$ represents an attenuated traveling wave propagating through the porous medium. The functional form $C_0(r_i(r,t))$ defines the wave shape, which is determined by the initial injection profile and translated radially according to the cubic trajectory $r(t)$. As the wave front advances, its amplitude undergoes a homogeneous exponential decay due to the filtration term. Unlike the variable porosity case, where the wave would also experience structural distortion (due to the $\varepsilon'(r)$ gradient term), here the wave maintains its relative profile, being scaled only by the constant filtration coefficient. This allows for a clear decoupling between the kinematics of the front (the cubic root dependence) and the dynamics of particle deposition (the exponential attenuation).

5 Conclusion

In this study, we developed a hierarchical analytical framework to investigate nanofluid transport in deformable porous media, providing a fundamental benchmark for the validation of future 2D and 3D numerical simulations within high-performance computing (HPC) environments. By comparing three distinct configurations, we have highlighted how the structural representation of the medium dictates the predicted efficiency of the injection process. The transition between the three considered configurations (fixed, linearized, and general porosity) reveals an increasing degree of physical coupling, which nevertheless corresponds to a rise in the analytical complexity of the proposed solutions, as summarized below:

- **Fixed Porosity Model (ε_0)**: Acting as a "zero-order" approximation, this case describes the transport as a standard attenuated traveling wave. While ideal for a rapid, order-of-magnitude estimation of the design parameters, it fails to capture the spatial redistribution of particles caused by the medium's heterogeneity.
- **Linearized Porosity Model**: This intermediate stage illustrates the fundamental interaction between the injection pressure p_i and the response of the porous medium. In this case, the porosity acts as a structural constraint that deforms the advancing concentration front. The width of the front no longer depends solely on the initial concentration but is modulated by the ratio between the porosity value at r and that at the injection point. This demonstrates that the 'mechanical control' of the injection can be strategically utilized to promote deeper particle penetration by locally modifying the pore space.
- **General Variable-Porosity Model ($\varepsilon(r)$)**: This formulation provides the most rigorous analytical reference. By accounting for the porosity gradient, it captures the local acceleration and deceleration of the fluid, offering a precise map of the deposition patterns. This model serves as the ultimate "gold standard" for calibrating complex numerical solvers.

Although preliminary, this work provides a strategic road map for future research. The exact solutions derived here allow for the calibration of numerical solvers before they are deployed to simulate large-scale, non-linear deformation regimes and asymmetric geometries. Moving toward 2D and 3D models will enable the exploration of medium anisotropy, using these analytical benchmarks to ensure physical consistency and mathematical convergence.

References

1. Agista, M.N., Andersen, P.Ø., Yu, Z.: Modelling nanofluid injection in porous media. J. Petrol. Sci. Eng. **176**, 836–849 (2019)
2. Di Michele, F., Pizzichelli, G., Mazzolai, B., Sinibaldi, E.: On the preliminary design of hyperthermia treatments based on infusion and heating of magnetic nanofluids. Math. Biosci. **262**, 105–116 (2015)
3. Elsayed, M.E.A., Ayoub, H.A., Helal, M.I.D., Sang, W., Shenef, Z., Abdelhafeez, I.A.: Nanotechnology-enabled soil management for sustainable agriculture: interactions, challenges, and prospects. Environ. Sci. Nano **12**, 2128–2153 (2025)
4. Krishnan, J., Shukla, S.: The behaviour of soil stabilised with nanoparticles: an extensive review of the present status and its applications. Arab. J. Geosci. **12**(436) (2019)
5. Li, Y., Zhou, J., Schneider, E., Xi, S.: A review on development of nanofluid preparation and characterization. Powder Technol. **196**(2), 89–101 (2009)
6. MacMinn, C.W., Dufresne, E.R., Wettlaufer, J.S.: Large deformations of a soft porous material. Phys. Rev. Appl. **5**(4), 044020 (2016)
7. Pizzichelli, G., Di Michele, F., Sinibaldi, E.: An analytical model for nanoparticles concentration resulting from infusion into Poroelastic brain tissue. Math. Biosci. **272**, 6–14 (2016)
8. Salemi, H., Iglauer, S., Rezagholilou, A., Sarmadivaleh, M.: Laboratory measurement of Biot's coefficient and pore pressure influence on Poroelastic rock behaviour. APPEA J. **58**, 182–189 (2018)
9. Sathish, N., Ahalya, M., Thirunavukkarasu, T.S., et al.: A comprehensive review on the novel approaches using nanomaterials for the remediation of soil and water pollution. Alex. Eng. J. **86**, 373–385 (2024)
10. Sharma, R., Kumar, V.: Nano enabled agriculture for sustainable soil. Waste Manag. Bull. **2**(1), 152–161 (2024)
11. Tong, Q., et al.: Research progress in nanofluid-enhanced oil recovery technology and mechanism. Molecules **28**(7478) (2023)
12. Yao, K., Habibian, M., O'Melia, C.: Water and waste water filtration. concepts and applications. Environ. Sci. Technol. **5**(11), 1105–1112 (1971)
13. Yusuf, E.O., Amber, I., Officer, S., Oluyemi, G.F.: Transport of nanoparticles in porous media and associated environmental impact: a review. J. Eng. Res. **12**(2), 275–284 (2024)
14. Zanbili, F., Poursattar Marjani, A.: Innovative green and bio-based approaches for photosensitive nanoparticle synthesis: a review on methodologies, characterization, and applications. Micro Nano Syst. Lett. **13**(1), 3 (2025)

Author Index

A
Akhtar, Sania 135
Aleskerov, Fuad 384
Alonso Ramos, Andrea 405
Amaral, Paula 259
Annunziata, Alfonso 270
Antoniou, Constantinos 300
Audu, Johnson Daddy 236

B
Baltow, Badee 135
Begic, Hamza 300
Blanco, Guillermo 173
Bogdanov, Alexander 327
Borruso, Giuseppe 311
Bruzzone, Francesco 217, 300
Bui, Thiloan 226
Buzrukova, Makhina 186

C
Calderón, Nuria Rodríguez 393
Capizzi, Patrizia 104
Carollo, Alessandra 104
Chiang, Yueh 327
Codroiu, Denis 351
Colica, Emanuele 54
Costea, Ilona 405
Cozzolino, Marilena 68

D
D'Amico, Sebastiano 54
De Cosmo, Francesco Pio 153
Del Giudice, Francesco Paolo 270
Deluka-Tibljaš, Aleksandra 338
Di Michele, Federica 247, 416
Dinu, Oana 351, 405
Do, Kienlam 226

E
Enomoto, Takeshi 117
Espín-López, P. 374

F
Fairag, Faisal 236
Fanfani, Marco 21
Farrugia, Mario 54
Flores, David Caballero 393

G
Galli, Federico 21
Galone, Luciano 54
Gao, Q. 374
Garau, Chiara 21
Gasparetto, Matilde 217
Gentile, Vincenzo 68
Greco, Michele 270
Guerrero-García, Pablo 197, 206

H
Hacar, Müslüm 365
Hacar, Özge Öztürk 365
Hasanov, Habibullo 186
Hendrix, Eligius M. T. 197, 206
Hermida-Ramon, Jose Manuel 291
Huda, Noor Ul 135
Hussain, Tahzeeb 393

I
Ilie, Anamaria 405

K
Khvatov, Valery 327
Klobučar, Mirna 338
Kosychev, Aleksei 384
Kouzeleas, Stelios 3

© 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 16766, pp. 427–428, 2027.
https://doi.org/10.1007/978-3-032-30539-8

L
Losada-Barreiro, Sonia 291
Luzi, G. 374

M
Majani, Marco Vincenzo 104
Mandado, Marcos 281
Mansoor, Muhammad Taha 135
Marin, Cristian 405
Martins, João 259
Martorana, Raffaele 104
Mauriello, Paolo 68
Melgarejo, Marco 173
Murgante, Beniamino 270

N
Nesi, Paolo 21
Nguyen, Vanmanh 226
Nocera, Silvio 217, 300
Nomura, Suzune 117

O
Otković, Irena Ištoka 338

P
Pala, Antonio 86
Palazuelos, Camilo 173
Pantaleo, Gianni 21
Pera, Donato 247, 416
Pérez, Alejandro 173
Perez-Gracia, Vega 393
Perretti, Biagio 270
Piroddi, Luca 54

R
Radu, Valentina 351
Ramos, Rui 259
Rehman, Sadaq-Ur 135
Reisch, Cordula 247
Rocha, Ana Maria A. C. 197, 206
Rosca, Eugen 351, 405
Rubino, Bruno 247, 416
Rusca, Aura 351, 405

S
Safarov, Bahodirhon 186
Salnikova, Elena 186
Sampalmieri, Rosella 416
Santamaria, Berta Viñes 393
Santos, David 259
Scorza, Francesco 270
Sgheri, Luca 153
Shchegoleva, Nadezhda 327
Siddiqui, Talha Ahmed 135
Simonyan, Ani 311
Skuric, Maja 351
Stere, Armand 351
Šurdonja, Sanja 338

T
Tran, Ductrung 226
Tran, Vanha 226
Tsolakidis, Ioannis 3

V
Vacca, Giuseppina 86
Vecchi, Enrica 37, 86
Vlachonasiou, Eleni 3

Z
Zucca, Marco 37