A Data-Driven Design Approach for Multi-scale Green Infrastructure Design: Integrating Landscape Connectivity and Phytoremediation in the Piedmont Region (IT)

Daniele Cannatella(✉), Max van der Waal, and Francesca Rizzetto

Delft University of Technology, 2628 CD Delft, The Netherlands
{d.cannatella,f.rizzetto}@tudelft.nl,
m.vanderwaal@student.tudelft.nl

Abstract. Rapid urbanization, intensive agriculture, and industrial activities in Europe have led to escalating landscape fragmentation and environmental pollution, posing threats to biodiversity, and causing degradation of vital natural resources over time. In response, the imperative for sustainable green infrastructure (GI) development has intensified to effectively mitigate the adverse impacts of human activities on landscapes and enhance their capacity to support life. However, despite the recognized importance of GI in addressing environmental challenges, its design and implementation remains a complex challenge, due to the difficulty to fully grasp and assess how GI provisions link to benefits at multiple scales, and how the latter can be maximized.

This paper introduces a data-driven approach for the integrated and multi-scalar design of GI, leveraging phytoremediation techniques to remediate polluted sites. By employing phytoremediation techniques at the local scale, this approach aims at enhancing local soil and water conditions while simultaneously reinforcing connectivity in regional ecological networks. The proposed approach initiates with a Morphological Spatial Pattern Analysis (MSPA) and Network Connectivity Analysis (MCA) to identify potential ecological sources and assess the impact of designed green corridors on the performance of existing regional networks. Subsequently, mapping and characterization of polluted sites are carried out in a GIS environment to identify strategic areas for the development of GI features at the local scale. This research makes use of the Novara area in the Piedmont region (IT) to test the above mentioned approach. Initial results show that quantifying and integrating landscape connectivity with the transformation of polluted sites is a viable way to optimize GI design allowing strategic design interventions on the regional and local scale. The proposed data-driven design approach offers a holistic solution to the challenges of urbanization and landscape fragmentation by providing a framework that enhances ecological resilience, optimizes resource allocation, and contributes to the creation of multifunctional landscapes.

Keywords: Green Infrastructure · Phytoremediation · Data-Driven Design · GIS

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
O. Gervasi et al. (Eds.): ICCSA 2024 Workshops, LNCS 14820, pp. 388–405, 2024.
https://doi.org/10.1007/978-3-031-65285-1_25

1 Introduction

The modern era witnesses an unprecedented surge in urbanization, intensified agricultural practices, and industrialization worldwide. This has resulted in the reshaping of natural landscapes and the alteration of their processes, jeopardizing their capacity to sustain life. Rapid transformation has broken up once continuous and healthy ecosystems into fragmented patches, endangering flora and fauna. Furthermore, industrial activities have resulted in contamination of vital resources such as soil, water, and air. The European continent boasts a rich industrial history that has left an indelible mark on its landscapes, caused by the production sector among others [1]. Recent estimates suggest that there are an estimated 2.5 million potentially polluted sites in Europe, with around 340,000 of them confirmed as contaminated and necessitating urgent remediation [2]. Nature-Based Solutions (NbS) are widely acknowledged as valuable approaches for remediating contaminated land. NbS make use of or mimic natural processes to address various environmental challenges, while providing the conditions to enhance biodiversity, mitigate pollution, improve climate resilience, and promote public health and well-being by integrating nature into urban and landscape planning [3, 4].

In particular, contaminated sites remediation can leverage phytoremediation techniques to harness vegetation to extract, sequester, or degrade contaminants in soil and water. Phytoremediation can deliver a range of benefits that go beyond mere pollution mitigation. Moreover, as phytoremediation techniques rely on natural processes to remediate polluted soil and water, they prove more cost-effective than traditional engineering approaches. When integrated into broader spatial design strategies, it becomes a powerful tool for ecological restoration, strengthening degraded ecosystems and enhancing environmental quality and biodiversity. In this regard, wasted land redevelopment through phytoremediation can become a platform for crafting spatial strategies that advance sustainable development addressing multiple spatial scales, akin to the principles of Green Infrastructure (GI) planning and design.

GI are interconnected networks of natural and semi-natural spatial features that provide social, economic, and environmental benefits in urban and rural settings by integrating conservation with land development and infrastructure planning [5, 6] GI planning and design bring together multiple spatial scales, ranging from the small, fragmented system of open spaces in dense urban environments to larger regional scales, tailored to address the unique challenges and characteristics of each context [7]. Embracing this multi-scale approach ensures the safeguarding of pivotal natural areas like forests and wetlands, while also bolstering the ecological integrity of broader networks and simultaneously supporting local needs such as urban, social and economic revitalization of discarded areas.

Nonetheless, limitations in GI design and planning persist, including a tendency to undervalue ecosystem services and overlook territorial perspectives [8], which can hinder effective implementation at regional scales. Challenges also arise from insufficient integration into broader spatial plans [9] and the difficulty of adopting effective multidisciplinary approaches [10]. Additionally, overlooking existing spatial tools can lead to the design and development of green corridors that fail to address all biodiversity requirements [11]. To overcome these hurdles, leveraging available data and digital technologies becomes essential. Such tools empower designers and planners to rigorously

assess the performance of proposed GI interventions and optimize their impacts across multiple scales.

This paper introduces a novel data-driven spatial design approach to integrate and address multiple scales in GI planning. It highlights the use of phytoremediation techniques to remediate contaminated land as part of this approach. By identifying optimal ecological corridors at the regional scale and different typologies of contaminated areas, it pinpoints potential transformative pathways at the local scale, to improve soil and water quality while also bolstering connectivity within broader regional ecological networks.

This contribution is developed in five distinctive sections. Starting with Sect. 2, where the theoretical framework is explained towards data-driven spatial design, green infrastructure, landscape ecology and phytoremediation approach. Subsequently Sect. 3 introduces the methodological approach developed and further its application in the territory of Novara province, Piedmont Region (IT). In here the emphasis is given to the significance of quantifying landscape connectivity in designing for the optimization of GI and to delineate a methodology for the identification of contaminated sites typologies using k-means clustering method. Section 4 presents and discusses the outcomes of the analysis. Ultimately, Sect. 5 synthesizes the findings, highlighting the potential offered by the proposed data-driven design approach. This section addresses, also, the challenges related to urbanization and landscape fragmentation, while underlying the need for ongoing data quality assessment and refinement for practical application.

2 Theoretical Background

2.1 Data-Driven Spatial Design

Data-Driven Spatial Design aims to promote long-term sustainable spatial development by using digital technologies and computational workflows in combination with spatial design principles. It applies principles from urban informatics, science of cities, Geographic Information Science (GIS), and spatial design to spatially-oriented research, design and planning. In this way, Data-Driven Spatial Design can enhance decision-making in spatial design practice, supporting it with evidence-based methods for the planning, design and management of cities and territories, thus moving away from reliance on assumptions [12]. In spatial disciplines, Data-Driven Design provides designers with tools to generate new, systematic, and robust knowledge, offering fresh insights into the functioning of urban and natural systems. By analyzing data collected across various spatial and temporal scales using computer algorithms, this approach identifies social and ecological patterns and trends, unveiling spatial relationships and phenomena [13].

Data-Driven design bridges the gap between research and design, integrating digital tools with more ‘traditional’ methods available to designer. Digital technology amplifies the cognitive abilities of designers by providing computer-based tools and methods, thereby enriching their toolbox for addressing complex spatial challenges like GI planning and design. Throughout such a process, the designer is in charge of selecting the most suitable methods and tools for acquiring data and processing information, synthesizing it into new knowledge and visualizing it across different media, including static and dynamic mappings, plans, 3D models, and beyond.

When integrated into GI planning and design, data-driven approaches bring about evidence on environmental conditions, including climate patterns, hydrological dynamics, biodiversity and land use, among others. This enables designers to swiftly pinpoint areas requiring green infrastructure intervention. Such approaches facilitate the optimization of time, economic efforts, and resources by identifying the most effective solutions for specific environmental challenges and prioritizing projects to deliver multiple benefits to both environment and society. Additionally, they establish frameworks for monitoring the performance of GI over time and evaluating their efficacy. Moreover, as GI can support climate change impacts adaptation, they can support smarter, faster, and more holistic adaptive solutions [14] by offering standardized workflows that harness the wide range of openly available environmental data from European platforms like Copernicus, the European Union's Earth Observation Program, integrating it with national, regional and local datasets. Finally, Data-Driven Design Approaches aspire toward reproducible, scalable, and automated workflows, fostering clearer and more transparent decision-making processes and promoting replicability in local contexts.

2.2 Landscape Ecology and the Role of Green Infrastructure

To properly address the redesign of existing green infrastructure, this proposed methodology draws from landscape ecology theory and principles [15]. Landscape ecology focuses on the study of ecological patterns and processes within structured mosaics of adjacent, heterogeneous ecosystems [16]. It categorizes landscape features into three primary components: patches, corridors, and matrix. Patches are defined as separate areas of natural vegetation, corridors denote continuous connections between patches, and the matrix encompasses the remaining existing mixture of landscape components [15].

Landscape Ecology principles have been extensively integrated into the assessment of ecological networks and connectivity, contributing to the understanding of ecosystems' qualities like their composition and configuration [17], and their capacity to deliver ecosystem services across urban and natural landscapes [18]. It emphasizes the role and the significance of ecological networks in achieving landscape sustainability, as more connected networks better sustain life and biodiversity [19–24]. Ecological network connectivity is an essential aspect to be taken into account for wildlife conservation, serving as a foundational element for conservation planning and landscape management. The value of connectivity is highlighted in the context of landscape structure, where it is quantified as either structural connectivity (the pattern of the network in a given space) or functional connectivity (how a network is actually used) [18]. Therefore, designing green corridors to enhance connectivity is one of the main objectives in GI planning. Designing green corridors to bolster connectivity in spatial design traditionally involves visually exploring landscape patterns across various thematic layers on a map. These layers depict physical and functional aspects of a landscape, such as morphology, water system, protected ecological or high biodiversity areas, mobility infrastructures, land use, among others. By identifying and qualitatively evaluating the impact of spatial opportunities to strengthen corridors and connect existing ecological patches, designers can propose transformative actions for underused or abandoned areas. They can also pinpoint natural or man-made features hindering animal movement along ecological

corridors, like mobility infrastructure. While visually investigating landscape patterns is a valuable method, it lacks the ability to quantify the impacts of proposed design alternatives. GIS-based methods and tools enable the measurement of landscape connectivity and support the assessment of proposed spatial configurations by integrating a quantitative component into the design process. This quantitative analysis guides designers in decision-making by assessing the current conditions of a landscape and possible spatial configuration to optimize the connectivity of ecological networks and identify strategic areas for redevelopment.

2.3 Phytoremediation and Adaptive Transformation

Environmental pollution caused by human activities is impacting the quality of our living environment. In recent years, both in the urban and agriculture environment there has been an increasing recognition of the critical importance of soil fulfilling its natural functions [25]. The concerns about healthy soil have been caused by diverse factors such as accelerated climate change, geopolitical tensions, and the constraints imposed by planetary boundaries, such as resource scarcity, biodiversity loss, desertification, and pollution [26]. Phytoremediation related to soil is considered a valuable contribution for a clean healthy soil. It is often studied from the specific perspective of soil remediation and mentioned as a nature-based alternative to traditional chemical and mechanical solutions [27]. As an example, in 1997 a new masterplan for Westergasfabriek in the Amsterdam west area was designed following these principles. Central to the approach was the necessity to treat the contaminated soil on site instead of transporting it offsite. A cut-and-fill balance was planned, expecting new added cleaned soil after displacing polluted soil. The polluted soil was then reused to reshape the area creating a new system of small hills where the phytoremediation process was taking place. Instead, all around the buildings the existing ground levels were maintained. Nowadays the area serves as a multifunctional public space where some of the former industrial buildings host event and entertainment places. The large park is a relief for the densely inhabited neighborhoods in the surrounding areas [28]. In general phytoremediation offers several advantages over other remediation strategies: low cost and limited maintenance, zero impact on the environment due to the nature of the process, and large public acceptance as an attractive green technology [29].

Being exposed to a variety of natural allelochemicals and xenobiotics, plants have developed diverse detoxification mechanisms of organic compounds [30] to thrive in conditions altered by intense anthropic activity [31] and alleviate pollution. In addition, phytoremediation offers potential beneficial side-effects, such as erosion control, site restoration, carbon sequestration, and feedstock for biofuel production [31, 32]. It can contribute to increasing the biodiversity in the area subject to remediation, such as the case of the Landschaftspark in Duisburg-Nord, Germany. When the former steel works closed in 1985, the site was left highly polluted specially by heavy metals. Although the project did not address phytoremediation as the main project's strategy, it is an interesting example of how phytoremediation can be injected in the design process as a strategy to effectively provide new healthy local ecologies and subsequently more attractive public amenities. A variety of treatments to extract and secure contamination were part of the design and nowadays the park has become a refuge for rare and endangered plants,

while a new ecology is returning to the site [34]. By means of phytoremediation, a chain of subsequent activities may arise and set the conditions for a livable and sustainable city. Moreover, as the remediation process reduces the concentrations of contaminants, an increase of soil prices is often likely to be expected [27]. The Park de Ceuvel in Amsterdam [35] is a small-scale intervention on a brownfield, where the soil remains untouched while cleaned up with Phyto-solutions, and the implementation of active programs revitalized the area. The area has been giving place to valuable workspace, hospitality, meeting places and a green outdoor space. After the opening of De Ceuvel, the surrounding neighborhood became more attractive and consequently housing prices arose.

During their process Phyto-plants have the capacity to uptake contaminants from the soil at a slow rate. Therefore, achieving a significant reduction in contaminant of soil through phytoremediation often necessitates more than a decade. This implies that the design must take into account long-term processes that take place in polluted sites and propose extra programs including temporary uses of the space to allow incremental added values in the area. In this perspective, plants play a crucial role in enhancing the quality of life for people and contrasting climate change, as larger plants and trees provide shade, which helps mitigate temperature increases during summertime, thereby fostering a more hospitable environment. All these physical interactions of plants with the environment improve the quality and livability for people, leading to a better social climate environment [27].

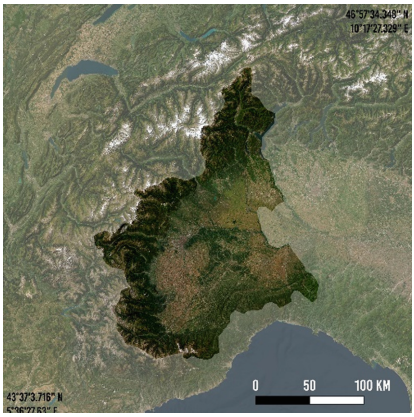


Fig. 1. Study area of Piemonte, IT

3 Materials and Methods

3.1 Study Area

The study area is located in the region between Turin and Milan (Italy), and spans approximately 14,000 km2 within the Piedmont administrative Region (Fig. 1). Bordered by the Alps to the northwestern side, the Turin hills to the south, and the Ticino River to the

east, it holds a crucial position in Italy's economic and industrial landscape. Indeed, this region lies within the historically renowned Italian Industrial Triangle, which contributed significantly to the growth of the country, accounting for 37.7% of the national industrial sector's value added [36]. The Po Valley, where the area is located, is also renowned for its fertile plains and temperate climate, serving as a vital agricultural hub within Italy and the EU. However, rapid urbanization has led to land-use conflicts resulting in habitat fragmentation and biodiversity loss [37, 38]. Moreover, contamination issues from intensive land use further exacerbate the challenges faced by the region. Given these complexities, the study area presents a compelling case for investigating strategies to enhance regional connectivity of the natural and semi-natural landscape through the redesign of GI in contaminated environments.

3.2 Data Sources

As this study aims to introduce a methodology that is highly replicable across diverse spatial contexts with similar landscape conditions in the EU, key to this approach is the use of accessible and open-source data and software. The primary software tools utilized include GeoDa 1.22[1], QGIS 3.23[2], GuidosToolbox 3.3[3], and Conefor 2.6[4].

Spatial data used for this study is sourced from authoritative repositories such as Copernicus Land Monitoring Service and open-access datasets sourced from the European Environment Agency (EEA), the Piedmont Region web-GIS platform, Open Street Map, and Google Maps/Street View. Table 1 provides an overview of the thematic layers used in the research, and their respective sources.

Table 1. Data sources overview

Theme	Thematic Layer	Source
Land-Use / Land Cover	CORINE Land Cover 2018 Vector	Copernicus Land Monitoring Service
Contaminated sites	Anagrafe di siti contaminati (ASCO)	Regione Piemonte - Servizi ambientali
	European Pollutant Release and Transfer Register (E-PRTR)	EEA
Site imagery	Google Street View Static API	Google
Street & railroad infrastructure	OpenStreetMap Overpass Turbo API	OpenStreetMap

(*continued*)

[1] https://geodacenter.github.io
[2] https://www.qgis.org
[3] https://forest.jrc.ec.europa.eu/en/activities/lpa/gtb/
[4] http://www.conefor.org

Table 1. (*continued*)

Theme	Thematic Layer	Source
Vegetation	Banca Dati Territoriale di Riferimento degli Enti (BDTRE) - Vegetazione	Regione Piemonte
Buildings & structure	Banca Dati Territoriale di Riferimento degli Enti (BDTRE) – Edifici e antropizzazione	Regione Piemonte

3.3 Regional Scale: Landscape Connectivity Analysis

The landscape connectivity of the study area at the regional scale was evaluated using two distinct approaches: Morphological Spatial Pattern Analysis (MSPA) and Graph-Theoretic Connectivity. The following paragraphs provide a concise introduction to each method.

Morphological Spatial Pattern Analysis (MSPA). Morphological Spatial Pattern Analysis (MSPA) serves as a valuable tool for classifying and segmenting binary patterns into seven distinctive categories. This raster-based method employs a single parameter to classify each pixel based on its spatial surroundings utilizing mathematical morphology operators [39]. These categories encompass cores, edges, perforations, bridges, loops, branches, and islets. The application of MSPA in spatial planning and landscape architecture aligns with Godron & Forman's landscape ecology theories [15], facilitating a comprehensive assessment of landscape structure. The adoption of MSPA has gained considerable attention in landscape conservation and planning studies, contributing to the establishment of landscape connectivity principles and the development of ecological network plans [40]. Furthermore, MSPA's scalability has led to its integration into GI planning at the urban scale, underscoring the significance of ecological connectivity in urban settings [41]. MSPA coupled with Geographic Information Systems (GIS) in planning ecological landscapes is substantial in GI planning and design. GIS provides a powerful platform for spatial analysis, enabling the integration of landscape metrics and the visualization of landscape patterns, such as the identification of spatial distribution of ecological features.

In this study, MSPA is used to identify ecological sources constituting the primary ecological structure of the region. The eight-neighborhood analysis method in Guidos Toolbox 3.3 software was employed for landscape pattern analysis. Given MSPA's suitability for analysis across various spatial scales, selecting fitting parameters was crucial for obtaining accurate results. Based on literature, the raster input resolution was set to 10mx10m, with an edge width of 5, resulting in 50m edge width. Core areas exceeding the threshold value of 8ha were chosen as ecological sources for connectivity analysis at the regional scale.

Graph-Theoretic Connectivity Analysis. Graph theory offers a robust framework for representing and analyzing landscape connectivity, offering insights into spatial relationships emerging from landscape structure and patterns [42]. While MSPA focuses

on the identification of patches, corridors, and other structural landscape components, the graph-theoretical network approach aims to highlight connections between these elements. Representing different components as nodes connected by links enables the visualization of a network on a map. Translating structural landscape components into a node-based network allows for more in-depth accurate calculations of connectivity between individual structural elements [42–44].

Conefor 2.6 software was used for Graph-Theoretic Connectivity Analysis (GTCA). Ecological sources previously identified in MSPA were imported as nodes, each carrying a unique identifier (ID) and their extent area. Conefor 2.6 software facilitates two types of landscape connectivity analyses, providing several indexes. This study primarily utilizes the Probability of Connectivity (PC) Index, defined as the probability that two randomly placed animals within the landscape fall into interconnected habitat areas, to calculate landscape connectivity. PC is non-binary and indicates lower connectivity whenever the distance between patches increases [45]. Connectivity between two sources is represented as the probability that species would travel from one source to another based on mean dispersion values.

To calculate connectivity with Conefor 2.6, several input parameters where used. The mean dispersion distance of the Western Honeybee (Apis mellifera L.) was used as representative value for pollinating species within the region, based on Hagler et al. [46]. Moreover, the Link Improvement-function in Conefor 2.6 allowed the assessment of potential connections between ecological sources. The impact of potential connections was quantified through dPC connector values, with higher values representing potential green corridors with a greater impact on overall landscape connectivity, prioritized in GI planning and design.

3.4 Local Scale: Mapping and Characterization of Polluted Sites

Mapping Contaminated Sites and Typology Analysis. In order to incorporate contaminated sites as key elements in the redesign of green infrastructure (GI), it was essential to comprehend the spatial distribution and conditions of these sites. To achieve this, various sources such as the European Pollutant Release and Transfer Register (E-PRTR) and the Anagrafe Regionale dei Siti Contaminati (ASCO) of Piedmont Region were collated and standardized within the QGIS environment. This process aimed to create a comprehensive and cohesive map of contaminated sites in the region. However, the available information on pollution sources, intensity, types of pollutants, and site conditions was often incomplete.

Nonetheless, a geodatabase was developed, containing information about the geometrical and spatial information for each polluted site, as well as information on size, former use, type of pollutant and polluted media. This information was complemented with information on the surrounding predominant landscape. For missing data on current use of the sites, a neural network was established to approximate the missing data on operability. Using QGIS and Dutch Cadastral data, which does contain detailed information on operability of building, the coordinates of 7186 inoperable and 7501 inoperable buildings were extracted. Google's Street View Static API and Python, images from Google Street View were downloaded of these buildings. As not all of these buildings

had adequate imagery, this resulted in 4,340 images of operable buildings and 2,345 images of inoperable buildings. A neural network based on TensorFlow was developed, focused on image classification and trained on the extracted Google Street View imagery. Using the same approach, imagery for the contaminated sites within the research area were extracted. Using the neural network, these sites were classified as either operable or inoperable. This data was added to the existing geodatabase.

To characterize contaminated sites, a K-means cluster analysis was conducted to identify clusters in GeoDa. These clusters corresponded to typologies of the contaminated sites. Following the elbow-method, an optimal k-value of 5 clusters was determined.

Phytotechnologies and Design Solutions Suitability Assessment. Literature review and design references were used to design spatial interventions on the local scale. These interventions are linked to the aforementioned contamination site typologies based on spatial configuration and suitability within the spatial context. Suitability for contamination type and concentration, applicability in combination with soil characteristics and climate aspects are not taken into account yet. Detailed implementation of the spatial design interventions, including the plant species for phytoremediation, phasing based on contaminant types, maintenance regime, etc. are carried out for the specific characteristics of the individual site that is remediated.

4 Results and Discussion

4.1 Landscape Connectivity Analysis

The MSPA results revealed the ecological landscape composition of the Piedmont region, identifying 1699 ecological core areas, which account for 16,60% of the entire landscape (Fig. 2). The majority of these core areas are notably large, with sizes exceeding 40 ha constituting 15,81% of the overall landscape. These substantial ecological sources are mainly located along the foothills of the Alps, the Turin hills in the south, and the riparian green structures along the Po River and its main tributaries.

There are 37 cores with a size between 8ha and 40ha in the research area, comprising 0,48% of the total landscape. The ecological cores are mainly located in the plains nestled between mountainous regions; most of the former are located near the larger ecological cores of Piedmont region. Concentration of ecological cores decreases when the distance to the large ecological cores increases.

A total of 203 ecological cores with an area larger than 8 ha were selected for the connectivity analysis in Conefor. Table 2 provides an overview of the categorized ecological features.

The assessment of the ecological connectivity results in overall landscape connectivity values of 287.25x10^9 for PC and 53.5961.5 for the Equivalent Connected Area (ECA(PC)). The map in Figure X offers a graphical depiction of the link enhancement analysis performed in Conefor 2.6. It illustrates the effects of reinforcing existing connections and introducing new ones between habitat patches, with a color gradient ranging from blue (representing minimal impact) to red (signifying high impact). Due to computational limitations and processing considerations, a threshold of maximum length

20km was imposed. Additionally, a lower threshold of 5 km was applied, as corridors shorter that this were considered unsuitable for regional-scale development projects. Results show that several areas with a notable concentration of potential links are evident across the study area. Predominantly, these hotspots are situated along the foothills of the Alps, the Turin Hills, and the region between the Po River and its primary tributaries. Notably, the southeastern sector, encompassing the area between Mortara and Pavia, exhibits a high density of promising links. However, it was excluded because it lacks sites requiring remediation. On the other hand, the area extending from Chivasso to the foothills of the Alps showcases both high-potential links and a significant presence of wasteland. Yet, the monotonous and densely industrialized landscape was deemed unattractive from a design standpoint. An advantageous balance between contaminated sites and high-impact links is observed in the area encompassing Vercelli, Novara, and the Ticino Valley Natural Park. Here there is a high density of contaminated sites dispersed throughout a varied landscape including agricultural areas, riparian zones, urban centers, and industrial zones. The selected corridor linking Vercelli and the Ticino Valley Natural Park boasts a dPC value of 0.34, positioning it in the top 4% of potential link enhancement. Additional connections between Vercelli and Novara and the Ticino River contribute further to the enhancement of landscape values.



Fig. 2. MSPA results

Table 2. Statistics on landscape composition from MSPA

Component	Ratio (Core/Area)	Count
CORE(s) [green]	1.38/0.31	1466
CORE(m) [green]	2.11/0.48	135
CORE(l) [green]	69.52/15.81	68
ISLET [brown]	0.09/0.02	31

(*continued*)

Table 2. (*continued*)

Component	Ratio (Core/Area)	Count
PERFORATION [blue]	2.79/0.63	198
EDGE [black]	19.19/4.36	1046
LOOP [yellow]	0.11/0.03	158
BRIDGE [red]	0.99/0.23	1246
BRANCH [orange]	3.80/0.87	21633
Background [grey]	--/77.26	505/37158230

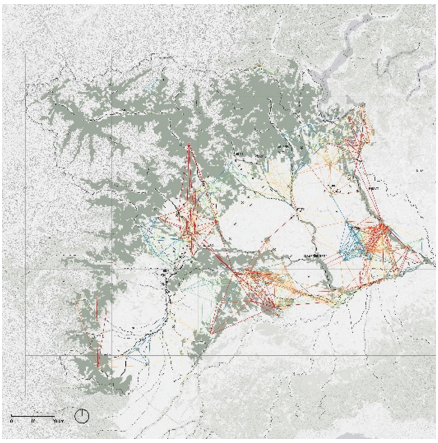


Fig. 3. Results from Link Improvement analysis in Conefor 2.6. Showing dPCconnector values.

4.2 Contaminated Site Typologies

To understand the spatial distribution of polluted areas, 171 of the contaminated sites within the previously selected corridor area were mapped and researched. Available data included plot size, presence of buildings and/or structures, surrounding context and land use, and site operability. Regrettably, operational status data for contaminated sites were not always readily accessible. To address this gap, a Convolutional Neural Network (CNN) was employed for image recognition. Using the images sourced from Google Street View, the operability status of the sites was determined, and results were subsequently added to the geodatabase to perform the cluster analysis in GeoDa to identify five typologies of contaminated sites.

K-means cluster analysis in GeoDa resulted in five clusters (Table 3). Using these typologies, design strategies and interventions can be established to redevelop these sites into more sustainable locations and make them integral parts of the GI planning and design process using phytoremediation and other nature-based solutions.

Typologies. Interpreting the results from the cluster analysis leads to 5 different typologies of contaminated sites. These are the typologies:

Table 3. Results from k-means clustering of the 83 remediation sites.

Cluster 1	*Mean*	*Median*	*Mode*
Structure [y/n]	*0*	*0*	*0*
Plot size [sqm]	*2404*	*1093*	*-*
*Land cover**	*2.2*	*2*	*2*
Operability [y/n]	*0.53*	*1*	*1*

Cluster 2	*Mean*	*Median*	*Mode*
Structure [y/n]	*1*	*1*	*1*
Plot size [sqm]	*4650*	*3529*	*-*
*Land cover**	*2.28*	*2*	*2*
Operability [y/n]	*0.64*	*1*	*1*

Cluster 3	*Mean*	*Median*	*Mode*
Structure [y/n]	*0*	*0*	*0*
Plot size [sqm]	*2404*	*1093*	*-*
*Land cover**	*2.2*	*2*	*2*
Operability [y/n]	*0.53*	*1*	*1*

Cluster 4	*Mean*	*Median*	*Mode*
Structure [y/n]	*1*	*1*	*1*
Plot size [sqm]	*4650*	*3529*	*-*
*Land cover**	*2.28*	*2*	*2*
Operability [y/n]	*0.64*	*1*	*1*

Cluster 5	*Mean*	*Median*	*Mode*
Structure [y/n]	*0*	*0*	*0*
Plot size [sqm]	*2404*	*1093*	*-*
*Land cover**	*2.2*	*2*	*2*
Operability [y/n]	*0.53*	*1*	*1*

**For land cover applies: 0 = nature, 1 = agricultural, 2 = urban, 3 = industrial.*

- Cluster 1 - Urban Incidents: These sites are characterized by localized contamination resulting from incidents like spillage or pipeline bursts. Often small and devoid of buildings or other structures, and located in the immediate surroundings of urban areas, they may be fenced off. These sites are ideal for transformation into social programs or parks, combining urban functions with phytoremediation practices (Fig. 4).
- Cluster 2 - Small industries/Commercial areas: these sites are typically found in urban areas or toward the edges of industrial zones. Contamination arises from industrial activities such as spillage or mismanagement of waste or other products. The mean plot size of these features is relatively small (approximately 0.4 ha), and are often still active (Fig. 3). Redevelopment strategies may involve buffering the sites with vegetation, landscaping, and integrating green energy production with phytoremediation practices.

- Cluster 3 - Polluting agriculture: these are medium to large scale sites outside urban cores, contaminated with pesticides, organic pollutants, or heavy metals due to agricultural practices. Sites can sometimes present structures or buildings. Design strategies for this typology could encompass the reintroduction of sustainable agriculture by introducing tolerant crops together or in rotation with phytoremediation plant species, creating agritourism by redeveloping the existing structures or rewilding.
- Cluster 4 - Industrial giants: mega-sized heavy industry sites in industrial zones, often still active and polluted due to processing of harmful material. Redevelopment strategies for this typology may involve buffering from other (urban or agricultural) land uses, reafforestation, or development of green energy solutions.
- Cluster 5 - Peri-Urban Plains: these are large contaminated plots due to former activity, often found at the edge of the urban cores, like the Silvio Piola Sports Park in Novara or the former airport of Vercelli. Areas within this typology typically do not have any structures, and are completely abandoned and gated to prevent human interaction with the contaminants. As these locations often form the border between urban areas and the agricultural landscape, these areas can be redeveloped into peri-urban parks or used for rural-urban integration with orchards or community gardens.



Fig. 4. Example sites of Cluster 1 – Urban Incident (Left) and Cluster 2 – Small Industries (right)

5 Conclusions

The implementation of a data/driven design approach in GI planning and design within urban, peri-urban and rural landscapes, as demonstrated in this research, offers significant potential to improve the efficiency and the effectiveness of the design process, bridging together different spatial scales and informing strategic decision-making in the design process. By synthesizing existing literature into a comprehensive approach and using different digital tools, this study showcased the role of data as a primary design input at both regional and local scales. The identification of potential green corridors in the Piedmont Region, together with the classification of contaminated sites into distinct typologies, provides valuable insights for devising site-specific design interventions and strategies that can generate benefits on larger scale networks in terms of ecological connectivity and biodiversity.

The integration of phytoremediation and nature-based solutions emerges as crucial elements in transforming contaminated sites into integral components of the regional green infrastructure network in the study area, where multiple potential sites for GI-design were identified. MSPA showed 203 ecological cores with a relevant area size of over 8 ha. While the largest ecological sources are situated in the mountainous and hilly areas of the Alps and the Turin hills, and in the form of riparian zones along the main rivers, a great number of ecological sources were found in the plain. Overlaying the identified ecological sources with the mapped contaminated sites allowed the informed selection of the design area. After conducting a connectivity analysis, the area between the cities of Vercelli, Novara and the Ticino Valley Natural Park was selected as the location with a high potential impact on ecological connectivity. This area was chosen because of the high concentration of contaminated sites and a dynamic landscape consisting of urban cores, agricultural landscapes and industrial zones.

On the local scale, a data-driven analysis was used to identify different typologies of contaminated sites and classify the sites accordingly. A k-means cluster analysis based on relevant criteria was conducted in GeoDa. The resulting clusters were based on plot size, presence of buildings/structures on the site, operability of the sites, and the (surrounding) land cover. This data was gathered from the Corine Land Cover-dataset, cadastral data, OpenStreetMap and Google Street View-imagery, ensuring widespread accessibility and availability for different application locations.

The outcomes of the k-means analysis revealed five distinct typologies of contaminated sites: A. Urban Incidents, B. Small industries, C. Polluting farms, D. Industrial giants, E. Peri-urban plains. For each of the abovementioned typologies, several design interventions and strategies were identified and served as foundational constructs for devising site-specific design principles that help integrate contaminated sites into a broader green corridor redesign strategy.

However, the efficacy of this approach is contingent upon the accuracy and the availability of input data. Challenges such as incomplete or missing data, lack of homogeneity of datasets from different agencies or administrative entities, and biases in training and input data underscore the importance of thorough data auditing for the practical application of this approach. For instance, out of the 171 sampled contaminated locations, only 83 had accessible imagery, resulting in a significant decrease in the sample size for the CNN. Additionally, assessing contamination extent relies on cadastral data for plot size, yet it's crucial to recognize that contamination can extend beyond administrative boundaries, potentially affecting areas outside the delineated plots. Despite these challenges, the data-driven design approach holds promise for guiding strategic decision-making processes in GI planning and design.

Given its data-driven nature, the efficacy and precision of this approach, along with the subsequent formulation of design strategies and decisions, are highly dependent on the accuracy of the gathered input data. Throughout the course of this case study, the presence of deficient or erroneous data has caused inconsistencies within the analytical outcomes. E.g. among the 171 sampled contaminated locations, merely 83 possessed available imagery, meaning a substantial reduction in the sample pool. Furthermore, in assessing the extent of contamination, reliance on plot size derived from cadastral data is necessitated. Nonetheless, it is imperative to acknowledge that contamination is not

bound by administrative boundaries, thereby potentially manifesting in areas beyond the delineated plot area. Additionally, certain contaminated sites are situated within public spaces, devoid of definitive cadastral boundaries, resulting in null values for plot size. Compounding these challenges, both training and input data utilized for analysis with the CNN frequently exhibit biases, censorship, misalignment, or alterations, thereby exacerbating the complexity of operability assessment via image recognition. Many instances of erroneous representations within images, including misidentification of structures, delineation of merely the surrounding public space bordering the actual site, or presence of blurred regions due to stitching or censorship in Google Street View, further decrease the accuracy of the analysis. A comprehensive data audit is an important prerequisite for this approach to become applicable in practice.

Disclosure of Interests. The authors declare no conflict of interest.

References

1. Paya Perez, A., Rodriguez Eugeniom, N.: Status of local soil contamination in Europe: revision of the indicator "progress in the management of contaminated sites in Europe", publication office of the Europeans union. Luxembourg (2018). https://doi.org/10.2760/093804
2. Panagos, P., Van Liedekerke, M., Yigini, Y., Montanarella, L.: Contaminated sites in Europe: review of the current situation based on data collected through a European network. J. Environ. Public Health (2013). https://doi.org/10.1155/2013/158764
3. Zhao, S., Yue, B.: Nature-based solutions: establishing a comprehensive framework for addressing urban waterlogging management. Integr. Environ. Assess. Manag. **19**, 1414–1421 (2023). https://doi.org/10.1002/ieam.4786
4. Randelović, A., Figueras, A., Seidelin, F., Briggs, L., Stanić, F.: Nature-based Solutions (NBS) at work and monitoring their performance – the innovative research case of the EU-funded project euPOLIS. In: EGU General Assembly, Vienna, Austria, 24–28 2023, EGU23–17063 (2023) https://doi.org/10.5194/egusphere-egu23-17063
5. Wirth, P., Chang, J., Syrbe, R.U., et al.: Green infrastructure: a planning concept for the urban transformation of former coal-mining cities. Int J Coal Sci Technol **5**, 78–91 (2018). https://doi.org/10.1007/s40789-018-0200-y
6. Benedict, M.A., McMahon, E.T.: Green infrastructure: smart conservation for the 21st century. **20**(3), 12–17 (2002)
7. Firehock, K., Walker, R.A: Strategic Green Infrastructure Planning: A Multi-Scale Approach. (2015) https://doi.org/10.5822/978-1-61091-693-6
8. Pozoukidou, G.: Designing a green infrastructure network for metropolitan areas: a spatial planning approach. Euro-Mediterr J Environ Integr **5**, 40 (2020). https://doi.org/10.1007/s41207-020-00178-8
9. Maryati, S., Humaira, A.N.S.: Implementation of green infrastructure concept in Citarum Watershed. In: AIP Conference Proceedings, **1818** (1), 020031 (2017) https://doi.org/10.1063/1.4976895
10. Sitzenfrei, R., Kleidorfer, M., Bach, P.M., Bacchin, T.K.: Green infrastructures for urban water system: balance between cities and nature. Water **12**, 1456 (2020). https://doi.org/10.3390/w12051456
11. Snäll, T., Lehtomäki, J., Arponen, A., et al.: Green infrastructure design based on spatial conservation prioritization and modeling of biodiversity features and ecosystem services. Environ. Manage. **57**, 251–256 (2016). https://doi.org/10.1007/s00267-015-0613-y

12. van Ameijde, J.: Data-driven urban design: conceptual and methodological constructs for people-oriented public spaces. In: Morel, P., Bier, H. (eds.) Disruptive Technologies: The convergence of new paradigms in architecture. Springer Series initiative in Adaptive Environments. Springer, Cham (2023)
13. NY Times. https://www.nytimes.com/2012/03/04/business/ibm-takes-smarter-cities-concept-to-rio-de-janeiro.html. Accessed 29 Apr 2024
14. European Commission: Forging a Climate-Resilient Europe. The new EU Strategy on Adaptation to Climate Change (2021)
15. Forman, R.T.T., Godron, M.: Landscape ecology. Wiley (1986)
16. Loreau, M., Mouquet, N., Holt, R.D.: Meta-ecosystems: a theoretical framework for a spatial ecosystem ecology. Ecol. Lett. **6**(8), 673–679 (2003). https://doi.org/10.1046/j.1461-0248.2003.00483.x
17. Cannatella D., Nijhuis, S.: Assessing urban landscape composition and configuration in the pearl river delta (China) over time. J. Digit. Landscape Architect. **5**, 111–121 (2020). https://doi.org/10.14627/537690012
18. Egerer, M., Anderson, E.G.: Social-ecological connectivity to understand ecosystem service provision across networks in urban landscapes. Land **9**(12), 530 (2020). https://doi.org/10.3390/land9120530
19. Taylor, P.D., Fahrig, L., Henein, K., Merriam, G.: Connectivity is a vital element of landscape structure. Oikos **68**(3), 571 (1993). https://doi.org/10.2307/3544927
20. Clergeau, P., Burel, F.: The role of spatio-temporal patch connectivity at the landscape level: an example in a bird distribution. Landscape Urban Plann. **38**(1–2), 37–43 (1997). https://doi.org/10.1016/s0169-2046(97)00017-0
21. With, K.A.: The application of neutral landscape models in conservation biology. Conserv. Biol. **11**(5), 1069–1080 (1997). https://doi.org/10.1046/j.1523-1739.1997.96210.x
22. Collinge, S.K.: Spatial arrangement of habitat patches and corridors: clues from ecological field experiments. Landscape Urban Plann. **42**(2–4), 157–168 (1998). https://doi.org/10.1016/s0169-2046(98)00085-1
23. Raison, R.J., Brown, A.G., Flinn, D.W.: Criteria and Indicators for Sustainable Forest Management. CABI (2001)
24. Crist, M.R., Wilmer, B., Aplet, G.H.: Assessing the value of roadless areas in a conservation reserve strategy: biodiversity and landscape connectivity in the northern rockies. J. Appl. Ecol. **42**(1), 181–191 (2005). https://doi.org/10.1111/j.1365-2664.2005.00996.x
25. Balestrini, R., Lumini E., Borriello, R., Bianciotto, V.: Plant-Soil Biota Interactions, Istituto per la Protezione Sostenibile delle Piante, UOS Torino, Viale Mattioli, 10125 Torino, Italy (chapter 11 in Soil Microbiology, Ecology and Biochemistry 4th Edition by Eldor Paul (Editor) (2013)
26. United Nations Environment Programme:. Making Peace with Nature: A scientific blueprint to tackle the climate, biodiversity and pollution emergencies. Nairobi. https://www.unep.org/resources/making-peace-nature ISBN: 978-92-807-3837-7 Job No: DEW/2335/NA, Report Leads: Ivar A. Baste (GEO, IPBES; Norwegian Environment Agency, Norway) and Robert T. Watson (IPCC, IPBES; UEA, UK) (2021)
27. Grotenhuis, T., van Vianen, L., Wijnja, A., Bruijnes, J.: Phytoremediation as Accelerator of Transformation towards Regenerative Cities, A Nature based solution for environmental pollution Report number 383 Wageningen,. ISBN 978–94–6447–465–7 https://doi.org/10.18174/579598(2022)
28. Gustafson, Porter + Bowman. https://www.gp-b.com/cultuurpark-westergasfabriek. Accessed 07 May 2024
29. Gerhardt, K.E., Huang, X.D., Glick, B.R., Greenberg, B.M.: Phytoremediation and rhizoremediation of organic soil contaminants: Potential and challenges. **176**, 20-30 (2009)

30. Singer, A.: The chemical ecology of pollutants biodegradation. In: Mackova, M., Dowling, D., Macek, T. (eds.) Phytoremediation and Rhizoremediation: Theoretical Background, pp. 5–21. Springer, Dordrecht (2006)
31. Van Aken, B., Lafferty Doty, S.: Transgenic plants and associated bacteria for phytoremediation of chlorinated compounds. Biotechnol. Genet. Eng. Rev. **26**(1), 43–64 (2009). https://doi.org/10.5661/bger-26-43
32. Dietz, A., Schnoor, J.: Advances in phytoremediation. Environ. Health Perspect. **109**, 163–168 (2001)
33. Doty, S.L., et al.: Enhanced phytoremediation of volatile environmental pollutants with transgenic trees. Proc Natl Acad Sci USA **104**, 16816-16821 (2007)
34. Latz und Partner. https://www.latzundpartner.de/en/projekte/postindustrielle-landschaften/landschaftspark-duisburg-nord-de/. Accessed 07 May 2024
35. Delva landscape. https://delva.la/projecten/de-ceuvel/ Accessed 07 May 2024
36. OECD. (2023). National Accounts regional main aggregates : Value added by industry. © OECD. http://dati.istat.it/Index.aspx?QueryId=11479&lang=en
37. Tscharntke, T., Klein, A., Kruess, A., Steffan-Dewenter, I., Thies, C.: Landscape perspectives on agricultural intensification and biodiversity – ecosystem service management. Ecol. Lett. **8**(8), 857–874 (2005). https://doi.org/10.1111/j.1461-0248.2005.00782.x
38. Grossauer, F., Stoeglehner, G.: Bioeconomy—spatial requirements for sustainable development. Sustainability **12**(5), 1877 (2020). https://doi.org/10.3390/su12051877
39. Soille, P., Vogt, P.: Morphological segmentation of binary patterns. Pattern Recogn. Lett. **30**(4), 456–459 (2009). https://doi.org/10.1016/j.patrec.2008.10.015
40. Jin, L., Xu, Q., Yi, J., Zhong, X.: Integrating CVOR-GWLR-Circuit model into construction of ecological security pattern in Yunnan Province, China. Environ. Sci. Pollut. Res. Int. **29**(54), 81520–81545 (2022). https://doi.org/10.1007/s11356-022-21421-5
41. Mei, Y., Sun, Y., Wang, Q., Liu, Q., Zhang, L.: Construction of green space ecological network in Jinan city based on MSPA and MCR model. Pol. J. Environ. Stud. **31**(4), 3701–3711 (2022). https://doi.org/10.15244/pjoes/146344
42. Bunn, A.G., Urban, D.L., Keitt, T.H.: Landscape connectivity: a conservation application of graph theory. J. Environ. Manag. **59**, 265–278 (2000)
43. Ricotta, C., Stanisci, A., Avena, G.C., Blasi, C.: Quantifying the network connectivity of landscape mosaics: a graph-theoretical approach. Commun. Ecol. **1**, 89–94 (2000)
44. Urban, D., Keitt, T.: Landscape connectivity: a graphtheoretic perspective. Ecology **82**, 1205–1218 (2001)
45. Saura, S., Pascual-Hortal, L.: A new habitat availability index to integrate connectivity in landscape conservation planning: comparison with existing indices and application to a case study. Landscape Urban Plann. **83**(2–3), 91–103 (2007). https://doi.org/10.1016/j.landurbplan.2007.03.005
46. Hagler, J.R., Mueller, S., Teuber, L.R., Machtley, S.A., Van Deynze, A.: Foraging range of honey bees, Apis mellifera, in alfalfa seed production fields. J. Insect Sci. **11**(144), 1–12 (2011). https://doi.org/10.1673/031.011.14401