Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

An Integrated Spatial Data Infrastructure for Sustainable Tourism Planning: Development and Implementation in the Lake Ohrid Region

Bashkim Idrizi
Temenoujka Bandrova
bandrova_fgs@uacg.bg

University of Architecture
Artan Rexhepi
University of Architecture

Research Article

Keywords: Spatial Data Infrastructure, tourism GIS, spatial modelling, sustainable tourism planning, multi-criteria evaluation, accessibility analysis, WebGIS, Lake Ohrid, tourism suitability analysis

Posted Date: October 15th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10220204/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Tourism planning increasingly requires integrated geospatial frameworks capable of supporting spatial analysis and sustainable decision-making. This study develops a tourism-oriented Spatial Data Infrastructure (SDI) designed to integrate heterogeneous spatial datasets and support spatial modelling for the Lake Ohrid region. The proposed approach combines structured geodatabase design with spatial analytical methods, including kernel density estimation, accessibility modelling, spatial autocorrelation analysis, and multi-criteria tourism suitability evaluation using an Analytic Hierarchy Process. The SDI integrates tourism infrastructure, transportation networks, cultural heritage resources, and environmental datasets within a unified geospatial database and WebGIS-based decision-support environment. Results reveal pronounced spatial clustering of tourism infrastructure along the southern lakeshore, with a majority of facilities located near the lake and in areas characterized by high accessibility and proximity to cultural heritage sites. The tourism suitability model identifies high-potential development zones that strongly correspond to the existing distribution of tourism infrastructure, and sensitivity analysis confirms the robustness of the modelling framework. The findings demonstrate that integrating SDI architecture with spatial modelling provides a reproducible geospatial decision-support framework for sustainable tourism planning that is transferable to other environmentally sensitive regions. The proposed framework illustrates a reproducible approach for integrating SDI architecture with spatial modelling in tourism planning and may be transferable to other environmentally sensitive regions.

Introduction

Tourism represents one of the fastest-growing sectors of the global economy and plays a significant role in regional development, employment generation, and cultural exchange (Hall and Page, 2014; Boers and Cottrell, 2007). Due to the inherently spatial nature of tourism activities, tourism systems are strongly influenced by geographic factors such as accessibility, environmental conditions, landscape attractiveness, and the distribution of cultural heritage resources (Hall and Page, 2014; Geneletti and Dawa, 2009). As a result, geospatial technologies have become essential tools for analyzing tourism dynamics and supporting evidence-based planning and decision-making processes (Bahaire and Elliott-White, 1999; McAdam, 1999).

Geographic Information Systems (GIS) have been widely applied in tourism research for tourism resource inventory, destination management, accessibility analysis, and spatial suitability modelling (Bahaire and Elliott-White, 1999; Farsari and Prastacos, 2004; Rahman et al., 2024). GIS-based approaches allow the integration and analysis of spatial relationships between tourism infrastructure, transportation networks, environmental conditions, and cultural heritage resources (Farsari and Prastacos, 2004). These analytical capabilities support the identification of suitable areas for tourism development and improve the understanding of spatial tourism patterns (Rahman et al., 2024). However, many tourism GIS applications remain focused on individual analyses or isolated datasets, often lacking a structured geospatial data management framework capable of integrating heterogeneous spatial information from multiple sources (Delgado Fernandez and Delgado Fernandez, 2012; Zhang and Li, 2005).

Spatial Data Infrastructures (SDIs) provide a comprehensive framework for the systematic organization, management, sharing, and dissemination of geospatial information (Rajabifard et al., 2002; Nebert, 2004; Crompvoets et al., 2004). SDIs facilitate interoperability, standardization, and accessibility of spatial data across institutions and application domains (INSPIRE, 2007; Bernard et al., 2005; Maguire and Longley, 2005). In tourism planning, Spatial Data Infrastructure (SDI) architectures can support the integration of tourism infrastructure, environmental indicators, cultural heritage resources, and socio-economic data within a unified geospatial environment (Delgado Fernandez and Delgado Fernandez, 2012; Budhathoki et al., 2008). Despite these advantages, the development of domain-specific SDIs tailored to tourism systems remains relatively limited.

Beyond their traditional role in data management and interoperability, modern SDIs are increasingly expected to support advanced spatial analytics, reproducible modelling workflows, and evidence-based decision-making. This evolution from passive geospatial repositories toward analytical geospatial platforms remains insufficiently explored within tourism-oriented SDI research (Budhathoki et al., 2008; Batty et al., 2010; Tsou, 2011).

Most existing tourism GIS studies focus primarily on spatial modelling techniques such as accessibility analysis, tourism suitability modelling, and resource mapping (Bahaire and Elliott-White, 1999; McAdam, 1999; Farsari and Prastacos, 2004). While these approaches provide valuable insights into tourism spatial patterns, they often rely on fragmented datasets that are not organized within structured geospatial infrastructures (Delgado Fernandez and Delgado Fernandez, 2012; Rajabifard et al., 2002). Consequently, the integration of heterogeneous datasets, the reproducibility of spatial modelling workflows, and the long-term maintenance of geospatial information remain challenging. These limitations reduce the transparency, consistency, and operational applicability of tourism spatial analyses, particularly in regions where decision-making requires the integration of environmental, cultural, infrastructural, and socio-economic information from multiple institutional sources. Furthermore, relatively few studies have explored how SDI architectures can be combined with advanced spatial modelling techniques and WebGIS technologies to support integrated tourism decision-support systems (Tsou, 2011; Zhao et al., 2012; Yao and Li, 2013).

This research addresses this gap by proposing an integrated Spatial Data Infrastructure designed specifically for tourism spatial analysis and sustainable tourism planning. The proposed framework combines geospatial database design, spatial statistical analysis, accessibility modelling, and multi-criteria spatial evaluation within a unified analytical environment. Rather than treating the Spatial Data Infrastructure solely as a repository for heterogeneous geospatial datasets, the proposed approach establishes an analytical framework in which data organization, spatial modelling, statistical analysis, and WebGIS-based visualization operate as interconnected components of a reproducible decision-support workflow. By integrating SDI principles with advanced spatial modelling techniques, the study demonstrates how domain-oriented geospatial infrastructures can support complex tourism planning processes through consistent data management, transparent analytical procedures, and evidence-based spatial decision-making (Budhathoki et al., 2008; Batty et al., 2010; Tsou, 2011).

We hypothesize that integrating SDI architecture with reproducible spatial modelling improves analytical consistency and supports more transparent tourism decision-making than conventional GIS-based workflows.

The methodological framework integrates heterogeneous spatial datasets related to tourism infrastructure, cultural heritage resources, environmental conditions, and transportation networks within a unified geospatial database (Longley et al., 2015). Spatial modelling techniques including kernel density estimation, spatial autocorrelation analysis, hotspot detection, accessibility modelling, and multi-criteria spatial evaluation are applied to analyze tourism spatial patterns and identify areas with high tourism development potential (Malczewski, 2006; Eastman, 2012).

The proposed approach is demonstrated through a case study of the Lake Ohrid region, one of the most important tourism destinations in Southeast Europe. Due to its exceptional natural and cultural values, the Lake Ohrid region has been designated as a UNESCO World Heritage site and represents an environmentally sensitive tourism destination requiring balanced development and conservation (Albrecht and Wilke, 2008; UNESCO, 2020; Geneletti and Dawa, 2009). This makes the region particularly suitable for evaluating the role of integrated geospatial infrastructures in supporting sustainable tourism planning.

Therefore, the primary objectives of this study are to: (i) develop a tourism-oriented Spatial Data Infrastructure integrating heterogeneous geospatial datasets within a unified database environment (Rajabifard et al., 2002; Nebert, 2004; Crompvoets et al., 2004); (ii) establish a reproducible spatial modelling framework combining spatial statistical analysis, accessibility modelling, and AHP-based multi-criteria evaluation (Longley et al., 2015; Malczewski, 2006; Eastman, 2012); and (iii) demonstrate how the integration of SDI architecture, spatial modelling, and WebGIS technologies (Peterson, 2008) can support evidence-based tourism planning and spatial decision-making in environmentally sensitive regions (Tsou, 2011; Zhao et al., 2012; Yao and Li, 2013).

The scientific contribution of this study is threefold. First, it proposes a domain-specific Spatial Data Infrastructure that extends the traditional role of SDIs from geospatial data repositories toward analytical platforms supporting tourism planning and spatial decision-making (Budhathoki et al., 2008; Batty et al., 2010; Tsou, 2011). Second, it develops a reproducible methodological framework that integrates geospatial database design, spatial statistical analysis, accessibility modelling, hotspot detection, and AHP-based multi-criteria evaluation (Izeiroski and Idrizi, 2018) within a unified analytical environment (Longley et al., 2015; Malczewski, 2006; Eastman, 2012). Third, the proposed framework is demonstrated through the Lake Ohrid region, illustrating how integrated SDI architectures and WebGIS technologies can support sustainable tourism planning and provide a transferable decision-support approach applicable to other environmentally sensitive tourism destinations (Maguire and Longley, 2005; Yao and Li, 2013; Gretzel et al., 2015).

Materials and methods

Study area

The study area covers the Lake Ohrid region, located in the southwestern part of North Macedonia and extending into Albania. Lake Ohrid is one of the oldest and deepest lakes in Europe, with an area of approximately 358 km² and a maximum depth of about 288 m. The region is characterized by high biodiversity, unique ecosystems, and significant natural values, making it an important tourism destination (Albrecht and Wilke, 2008).

The surrounding area includes the cities of Ohrid and Struga, as well as several smaller settlements along the lakeshore. The region contains diverse natural landscapes and numerous cultural and historical sites that support different forms of tourism, including cultural, heritage, and nature-based tourism (Hall and Page, 2014; Boers and Cottrell, 2007). The study area is characterized by pronounced spatial heterogeneity, including densely urbanized lakeshore zones, mountainous terrain, protected natural areas, dispersed rural settlements, and culturally significant sites, creating diverse spatial conditions for tourism development and geospatial analysis.

Due to its exceptional natural and cultural importance, the Lake Ohrid region has been designated as a UNESCO World Heritage site. The environmentally sensitive character of the area requires careful planning to balance tourism development with environmental protection, making the region suitable for spatial analysis and sustainable tourism planning (UNESCO, 2020; Geneletti and Dawa, 2009).

From a geospatial perspective, the Lake Ohrid region represents an appropriate case study because it combines high tourism intensity, heterogeneous spatial resources, complex environmental constraints, and cross-sectoral datasets that require coordinated geospatial management. These characteristics provide suitable conditions for evaluating the capability of an integrated Spatial Data Infrastructure to support multi-source data integration, spatial modelling, and evidence-based tourism planning within a geographically and environmentally sensitive region (UNESCO, 2020; Geneletti and Dawa, 2009).

Consequently, the Lake Ohrid region provides a representative environment for evaluating the proposed SDI-based analytical framework and assessing its applicability to sustainable tourism planning under complex spatial and environmental conditions.

Conceptual Framework for Tourism Spatial Data Infrastructure

Spatial Data Infrastructures (SDIs) provide a structured framework for the collection, management, analysis, and dissemination of geospatial data across multiple domains (Rajabifard et al., 2002; Nebert, 2004; Crompvoets et al., 2004). In tourism planning, SDI architectures enable the integration of heterogeneous spatial datasets, including tourism infrastructure, transportation networks, environmental indicators, and cultural heritage resources, within a unified geospatial environment (Delgado Fernandez and Delgado Fernandez, 2012; Maguire and Longley, 2005).

Unlike conventional Spatial Data Infrastructures primarily designed for data storage, interoperability, and dissemination, the proposed framework adopts an analytical perspective in which the SDI constitutes the operational backbone of the entire spatial modelling process. Within this architecture, geospatial data management, spatial analysis, statistical evaluation, and WebGIS visualization are integrated into a unified workflow that supports reproducible and evidence-based tourism planning.

The conceptual framework proposed in this study is based on a multi-layer SDI architecture that integrates data acquisition, geospatial database management, spatial analysis, and decision-support components. The data acquisition layer provides the fundamental geospatial input for the SDI by integrating heterogeneous datasets obtained from tourism inventories, transportation networks, environmental monitoring systems, and cultural heritage databases. Standardization at this stage ensures semantic consistency and interoperability across subsequent analytical procedures (Maguire and Longley, 2005). The geospatial database layer transforms heterogeneous datasets into a structured data environment through standardized feature classes, attribute models, and thematic organization. Beyond efficient storage, this layer establishes the data consistency required for reproducible spatial modelling and subsequent statistical analysis (Bernard et al., 2005; Longley et al., 2015).

The spatial analysis layer represents the analytical core of the framework by integrating kernel density estimation, spatial autocorrelation analysis, hotspot detection, accessibility modelling, and multi-criteria evaluation (Izeiroski et al., 2018). The combination of complementary analytical methods enables the identification of tourism spatial patterns from multiple methodological perspectives (Tsou, 2011; Malczewski, 2006).

The decision-support layer transforms analytical outputs into actionable geospatial information through interactive WebGIS visualization and spatial decision-support services. This layer facilitates communication of modelling results to planners and stakeholders while supporting transparent evidence-based tourism management (Zhao et al., 2012; Yao and Li, 2013).

The proposed architecture establishes a continuous analytical workflow in which each SDI layer directly supports the subsequent stage of the modelling process, from standardized data acquisition to spatial analysis and decision-support. Consequently, the Spatial Data Infrastructure functions not merely as a geospatial repository but as an integrated analytical platform enabling reproducible tourism spatial modelling and sustainable planning.

Conceptual framework of methodology

The methodological framework (Fig. 1) was designed as a sequential and reproducible geospatial workflow integrating data acquisition, geodatabase development, spatial modelling, spatial statistical analysis, and WebGIS-based decision support within a unified Spatial Data Infrastructure (Peterson, 2008). Rather than considering these components as independent analytical tasks, the proposed methodology treats them as interconnected stages of a continuous analytical workflow in which the output of each stage serves as the input for the subsequent analytical process.

The first stage involved the identification and collection of spatial datasets relevant to tourism planning. These datasets include tourism infrastructure, transportation networks, cultural heritage sites, administrative boundaries, land cover, and environmental indicators. The collected data were classified into thematic categories and standardized to ensure consistency in spatial resolution, coordinate systems, and attribute structure. Both vector and raster datasets were integrated into the geospatial database to support multi-scale spatial analysis. Before integration into the geodatabase, all datasets underwent harmonization procedures including coordinate reference system verification, attribute standardization, topology validation where applicable, and consistency checks to ensure interoperability between heterogeneous spatial data sources. These preprocessing procedures established the data quality required for subsequent spatial analyses and reproducible modelling.

In the second stage, the geodatabase was designed not only to facilitate efficient storage and management of spatial information but also to support analytical consistency throughout the modelling workflow. The hierarchical organization of thematic datasets, feature classes, and standardized attribute structures minimizes data redundancy, improves interoperability, and enables reproducible execution of successive spatial analytical procedures. The hierarchical organization of the Tourism SDI geodatabase is presented in Fig. 2, illustrating the relationships between thematic datasets, feature classes, and the analytical components that support subsequent spatial modelling procedures.

All spatial analyses were performed using open-source geospatial software to ensure methodological transparency and reproducibility. The analytical workflow combined QGIS for geospatial database management and visualization with Python-based geospatial libraries for spatial statistical analysis and automation of analytical procedures. WebGIS components (Haklay et al., 2008) were implemented using interoperable OGC-compliant services provided through GeoServer and QGIS Server.

The selected analytical methods were chosen because they represent complementary approaches for investigating tourism spatial systems. Kernel density estimation identifies spatial concentration patterns, Moran's I evaluates global spatial autocorrelation, Getis–Ord Gi detects statistically significant local clusters, accessibility analysis quantifies transportation connectivity (Gutierrez et al., 2008), while AHP-based multi-criteria evaluation (Idrizi et al., 2026) integrates heterogeneous spatial factors into a comprehensive tourism suitability model (Izeiroski and Idrizi, 2018). The combination of these methods enables tourism development to be assessed from distributional, statistical, network-based, and decision-support perspectives within a unified analytical framework.

Structure of the Spatial Data Infrastructure

The Spatial Data Infrastructure developed in this study is organized into eight thematic datasets (see Table 1) representing the main components of the regional tourism system (Eremchenko et al., 2016). These datasets integrate spatial information related to administrative boundaries, hydrography, transportation networks, settlements, cultural heritage, health and security facilities, tourism infrastructure, and environmental conditions.

The geospatial database contains 79 feature classes and 275 spatial feature types organized in a hierarchical structure. This organization enables efficient management and integration of heterogeneous spatial datasets and provides a comprehensive foundation for spatial analysis, tourism modelling, and decision-support within the proposed tourism Spatial Data Infrastructure.

Table 1
Structure of the geospatial database implemented in the tourism SDI.

Dataset	Description	No. of Feature Classes
Boundaries	Administrative and tourism zone boundaries	2
Hydrography	Lakes, rivers, wetlands, hydrological features	7
Transportation	Road network, airports, ports, parking	10
Settlements	Cities, villages, urban structures	5
Cultural heritage	Religious, historical and cultural sites	15
Health & security	Hospitals, police, emergency services	6
Tourism facilities	Hotels, restaurants, recreation	25
Environmental datasets	Climate and environmental indicators	9

Geodatabase Design

The geodatabase developed in this study constitutes the core component of the proposed Tourism Spatial Data Infrastructure, providing the structural foundation required for the integration, management, and analysis of heterogeneous geospatial information (Xiang and Fesenmaier, 2017). Following standard GIS database design principles, the database was developed through conceptual, logical, and physical modelling stages to ensure data consistency, interoperability, and reproducibility of subsequent spatial analyses (Longley et al., 2015).

Vector datasets represent discrete spatial entities describing the structural components of the tourism system, including settlements, transportation infrastructure, tourism facilities, administrative units, and cultural heritage resources. Their explicit geometric representation enables the analysis of spatial relationships, network connectivity, accessibility, and infrastructure distribution within the proposed analytical framework (Longley et al., 2015). Raster datasets represent spatially continuous environmental phenomena that influence tourism dynamics, including climatic variables, environmental conditions, and landscape characteristics. Their integration with vector datasets enables the simultaneous evaluation of discrete tourism infrastructure and continuous environmental gradients within a unified geospatial modelling environment. The integration of these datasets within a unified geodatabase enables multi-scale spatial analysis and supports tourism planning and spatial modelling.

The hierarchical organization of thematic datasets was intentionally designed to support successive analytical procedures presented in the following sections. Transportation and settlement layers provide the basis for accessibility modelling, tourism infrastructure layers support kernel density estimation and hotspot analysis, while environmental raster datasets contribute to multi-criteria tourism suitability evaluation. Consequently, the geodatabase functions not only as a storage environment but also as the operational structure supporting the complete analytical workflow. The conceptual database model was implemented as a structured GIS geodatabase consisting of thematic feature datasets organized according to the proposed SDI architecture. The implemented geodatabase structure is presented in Fig. 3.

A multi-criteria spatial evaluation model (Idrizi et al., 2026b) was implemented to identify areas with high tourism development potential. The model integrates spatial factors including accessibility to transportation networks, proximity to the lake shoreline, proximity to cultural heritage sites, landscape attractiveness, and environmental conditions. The conceptual workflow of the tourism suitability model is illustrated in Fig. 4. The Tourism Potential Index (TPI) was calculated using a weighted linear combination approach, with criterion weights derived using the Analytic Hierarchy Process (AHP; Malczewski, 2006; Eastman, 2012; Idrizi et al., 2026).

The developed geodatabase provides the common analytical foundation for all spatial modelling procedures implemented in this study. Rather than functioning solely as a repository of spatial information, each thematic dataset supports one or more analytical methods depending on its spatial characteristics and semantic content. The relationship between the principal thematic datasets, the corresponding analytical methods, and their expected outputs is summarized in Table 2. This overview highlights the logical integration between the proposed Spatial Data Infrastructure and the analytical workflow adopted for tourism planning.

Table 2
Relationship between thematic geospatial datasets, analytical methods, and expected outputs within the proposed Tourism Spatial Data Infrastructure.

Thematic dataset	Analytical method	Main analytical purpose	Output
Tourism facilities	Kernel Density Estimation (KDE)	Identify spatial concentration of tourism infrastructure	Tourism density surface
Tourism facilities	Moran's I Spatial Autocorrelation	Evaluate global spatial clustering	Moran's I statistics
Tourism facilities	Getis–Ord Gi* Hotspot Analysis	Detect statistically significant hotspots and cold spots	Hotspot map
Transportation network	Network accessibility analysis	Evaluate accessibility to tourism destinations	Accessibility map
Cultural heritage sites	Proximity analysis	Assess influence of cultural resources	Distance/proximity layers
Environmental raster datasets	Multi-criteria evaluation (AHP)	Quantify environmental suitability	Standardized suitability layers
Integrated thematic datasets	Weighted Linear Combination (WLC)	Calculate Tourism Potential Index (TPI)	Tourism suitability map
Complete SDI database	WebGIS visualization	Dissemination and decision support	Interactive WebGIS platform

All spatial criteria were standardized to a common scale to ensure comparability between datasets. The resulting suitability index was classified into tourism potential zones and integrated into the SDI environment. To evaluate the relationship between the suitability model and existing tourism development, a spatial correlation analysis was performed between tourism suitability values and tourism infrastructure density. The results indicate a strong positive correlation, confirming that the spatial modelling framework effectively captures the main drivers of tourism development. The spatial correlation between tourism suitability and tourism infrastructure distribution is presented in Fig. 5.

By integrating heterogeneous spatial datasets within a standardized hierarchical structure, the proposed geodatabase establishes the analytical foundation required for reproducible tourism spatial modelling, spatial statistical analysis, and WebGIS-based decision-support, thereby extending its function beyond conventional geospatial data management (Haklay, 2013).

RESULTS

Spatial Data Infrastructure Implementation

The developed Tourism Spatial Data Infrastructure successfully integrates heterogeneous geospatial datasets within a unified analytical environment comprising eight thematic datasets, 79 feature classes, and 275 spatial feature types. Beyond data organization, the implemented SDI establishes a standardized framework that enables consistent execution of spatial analytical procedures while maintaining interoperability among tourism infrastructure, transportation, environmental, and cultural heritage datasets. This integrated organization forms the analytical basis for all subsequent spatial modelling procedures presented in the following sections. To demonstrate the practical implementation of the proposed SDI, the integrated geospatial database and analytical results were published through an interactive WebGIS platform. The operational environment enables visualization of thematic datasets, tourism suitability results, and spatial query functions supporting tourism planning (Fig. 6).

The WebGIS implementation confirms that the proposed framework supports not only spatial data management but also interactive dissemination of analytical results for evidence-based tourism decision-making.

Spatial Distribution of Tourism Infrastructure

Kernel density estimation reveals a pronounced clustering of tourism infrastructure along the southern shoreline of Lake Ohrid. The highest density values exceed 45 facilities per km², particularly within the urban area of Ohrid and adjacent coastal zones.

Accessibility analysis indicates that approximately 72% of tourism facilities are located within 2 km of primary transportation routes, while nearly 85% are situated within 5 km of major cultural heritage sites. These results demonstrate that tourism infrastructure is strongly concentrated in areas characterized by high accessibility and proximity to key attractions. The observed concentration reflects the cumulative influence of lakeshore accessibility, urban development (Idrizi et al., 2018), and the spatial distribution of cultural attractions, indicating that tourism infrastructure is not randomly distributed but follows identifiable spatial development patterns.

Tourism Suitability Modelling

The multi-criteria tourism suitability model identifies the southern and southwestern lakeshore zones as areas with the highest tourism development potential. These areas combine favorable accessibility, lake proximity, landscape attractiveness, and cultural heritage availability (Sigala and Gretzel, 2018). Secondary suitability zones are observed near Struga and smaller lakeside settlements.

Comparison between model results and existing tourism infrastructure shows strong spatial agreement. Approximately 68% of tourism facilities are located within areas classified as high or very high suitability, indicating that the selected criteria effectively capture the main drivers of tourism development. These

findings indicate that transportation connectivity and proximity to major cultural attractions jointly represent the dominant spatial drivers influencing tourism infrastructure location within the study area.

Spatial Autocorrelation Analysis

Spatial autocorrelation analysis using Moran's I (Chem et al., 2010) reveals a statistically significant clustering pattern of tourism infrastructure. The calculated Moran's I value (I = 0.61) with a z-score greater than 2.5 and p-value below 0.01 confirms a strong positive spatial autocorrelation.

These results indicate that tourism infrastructure distribution is not random but spatially concentrated in lakeshore areas and zones with high accessibility and cultural heritage presence. The high spatial agreement between model predictions and the observed distribution of tourism facilities provides empirical support for the suitability model, suggesting that the selected evaluation criteria adequately represent the principal spatial determinants of tourism development in the Lake Ohrid region. The statistically significant positive spatial autocorrelation confirms that tourism infrastructure exhibits systematic clustering rather than random spatial distribution, supporting the subsequent hotspot analysis and validating the appropriateness of cluster-based spatial modelling approaches.

Hotspot Analysis

Hotspot analysis using the Getis–Ord Gi* statistic identifies statistically significant clusters of high tourism infrastructure density in the urban area of Ohrid and surrounding coastal zones. Cold spots are mainly located in inland mountainous areas characterized by limited accessibility and reduced tourism attractiveness.

The hotspot distribution confirms that tourism development follows a spatially clustered structure influenced by lake proximity and transportation connectivity. The coincidence between hotspot locations and areas identified by the tourism suitability model further demonstrates the internal consistency of the analytical framework and strengthens confidence in the integrated SDI-based modelling approach (Dissanayake et al., 2025).

Model Validation and Sensitivity Analysis

Spatial correlation analysis between tourism suitability and tourism infrastructure density shows a strong positive relationship ($r = 0.72$, $p < 0.01$). This result confirms (Chem et al., 2010) that the suitability model accurately represents the spatial drivers of tourism development.

Sensitivity analysis demonstrates model robustness under modified weighting scenarios. More than 85% of areas classified as high suitability in the baseline model remain stable across all tested configurations, with only minor variation in the extent of high-suitability zones. The limited variation observed under alternative weighting scenarios indicates that the proposed modelling framework is relatively insensitive to moderate changes in criterion importance, suggesting a stable analytical structure suitable for decision-support (Haklay, 2013) applications.

Summary of Spatial Modelling Results

The combined results of kernel density estimation, accessibility modelling, spatial autocorrelation, hotspot detection, and multi-criteria evaluation (Azizi et al., 2026) indicate that tourism development in the Lake Ohrid region exhibits a strongly clustered spatial structure. High tourism potential areas are primarily concentrated along the southern lakeshore, where accessibility conditions, landscape attractiveness, and cultural heritage resources overlap.

The developed SDI framework (Zwirowicz-Rutkowska and Michalik, 2024) enables the identification of priority tourism development zones and supports spatial decision-making for sustainable tourism planning (Haklay, 2013). Collectively, these findings demonstrate that the proposed Tourism Spatial Data Infrastructure functions as an integrated analytical environment in which standardized geospatial data management (Xiang and Fesenmaier, 2017), spatial statistical analysis (Idrizi et al., 2024), and multi-criteria evaluation (Izeiroski et al., 2018) jointly support transparent and reproducible tourism planning.

Discussion

The results demonstrate that integrating Spatial Data Infrastructure (Zwirowicz-Rutkowska and Michalik, 2024) with spatial modelling techniques (Xiang and Fesenmaier, 2017) provides a coherent analytical environment (Idrizi et al., 2024) for tourism planning. Rather than functioning solely as a geospatial data repository, the proposed SDI supports the systematic integration of heterogeneous datasets (Idrizi, 2006), spatial statistical analyses, and decision-support procedures within a unified and reproducible analytical workflow.

The observed spatial clustering of tourism infrastructure is consistent with previous studies reporting that tourism development is strongly influenced by accessibility, cultural heritage resources, and environmental attractiveness (Hall and Page, 2014; Geneletti and Dawa, 2009; Farsari and Prastacos, 2004). However, unlike most previous GIS-based tourism studies, the proposed framework integrates these analytical procedures within a unified Spatial Data Infrastructure (Zwirowicz-Rutkowska and Michalik, 2024), thereby improving data interoperability, methodological transparency, and reproducibility.

Spatial analysis revealed a pronounced clustering of tourism infrastructure along the southern shoreline of Lake Ohrid. Kernel density estimation, spatial autocorrelation, and hotspot analysis (Idrizi et al., 2026b; Idrizi et al., 2024) consistently indicated that tourism development is concentrated in lakeshore areas characterized by high accessibility and landscape attractiveness. These findings confirm that lake proximity and transportation connectivity (Gutierrez et al., 2008) are the dominant drivers of tourism spatial patterns.

The multi-criteria (Idrizi et al., 2026b) tourism suitability model identified priority development zones that largely correspond to the existing distribution of tourism infrastructure. The strong spatial agreement between suitability results and observed infrastructure confirms that the selected criteria effectively represent the main factors controlling tourism development. Sensitivity analysis (Azizi et al., 2026) further demonstrated the robustness of the model (Zwirowicz-Rutkowska and Michalik, 2024), showing stable suitability patterns under varying weighting scenarios.

Integrating the SDI framework with WebGIS (Haklay et al., 2008) visualization enhances the practical applicability of the approach by enabling interactive exploration of tourism datasets and modelling outputs. This supports evidence-based planning (Dodge and kitchin, 2007) and improves communication between stakeholders involved in tourism development.

Compared with conventional GIS-based tourism studies, where geospatial analyses are frequently performed as isolated workflows, the proposed framework integrates standardized data management, spatial modelling, statistical evaluation, and WebGIS dissemination within a single analytical environment. This integration improves methodological consistency, facilitates the reproducibility of spatial analyses, and reduces the fragmentation commonly associated with heterogeneous tourism datasets. Consequently, the proposed SDI architecture contributes not only to efficient geospatial data management but also to the operational implementation of transparent and evidence-based tourism planning.

Overall, the results confirm that SDI-based spatial modelling provides an effective decision-support framework for identifying suitable tourism development areas while considering accessibility, environmental constraints, and cultural heritage resources. Although the framework was developed using the Lake Ohrid region as a case study, its methodological structure is transferable to other tourism destinations where heterogeneous geospatial datasets, environmental constraints, and tourism infrastructure require integrated spatial analysis. Adaptation to other regions primarily requires the availability of standardized spatial datasets rather than modifications to the analytical workflow itself.

An important advantage of the proposed framework lies in the integration of database management and spatial modelling within a common analytical environment (Idrizi et al., 2018). This integration reduces inconsistencies arising from fragmented datasets, facilitates repeated analyses using standardized geospatial information, and improves the transparency of decision-support procedures.

Several limitations should be acknowledged. The analytical framework depends on the availability and quality of standardized spatial datasets, while the weighting scheme applied in the multi-criteria evaluation inevitably contains a degree of expert judgement (Izeiroski and Idrizi, 2018) despite the use of the Analytic Hierarchy Process (Idrizi et al., 2026b). Furthermore, tourism dynamics were analysed using predominantly static spatial datasets, whereas future developments could incorporate temporal mobility patterns, real-time tourism information, and continuously updated environmental observations to further improve decision-support capabilities.

In addition, the proposed framework was demonstrated using a single regional case study. Although the analytical workflow was intentionally designed to be transferable, its application to other tourism destinations with different environmental characteristics, governance structures, and data availability should be further investigated. Future comparative studies involving multiple case-study regions would provide additional evidence regarding the general applicability and scalability of the proposed SDI-based analytical framework.

Overall, the proposed Tourism Spatial Data Infrastructure demonstrates that combining standardized geospatial data management with advanced spatial analytical methods substantially enhances the capacity of geospatial decision-support systems for sustainable tourism planning. This confirms the value of integrating Spatial Data Infrastructures with analytical geospatial modelling for evidence-based tourism planning.

Conclusions

This study developed a tourism-oriented Spatial Data Infrastructure that integrates heterogeneous geospatial datasets, standardized geodatabase design, spatial statistical analysis, accessibility modelling, and multi-criteria evaluation within a unified analytical framework for sustainable tourism planning.

The implemented SDI, comprising eight thematic datasets, 79 feature classes, and 275 spatial feature types, provides the analytical foundation for reproducible spatial modelling and WebGIS-based decision support. The results demonstrate that tourism development in the Lake Ohrid region exhibits a pronounced spatial clustering controlled primarily by accessibility, lakeshore proximity, and cultural heritage resources, while the proposed suitability model successfully identifies priority development zones consistent with the observed distribution of tourism infrastructure.

Beyond the presented case study, the principal contribution of this research lies in demonstrating how a domain-specific Spatial Data Infrastructure can function as an integrated analytical platform that connects geospatial data management, spatial modelling, and interactive decision support within a coherent and reproducible workflow.

The proposed framework is transferable to other tourism destinations where heterogeneous geospatial datasets require integrated analysis for sustainable planning. Future research should focus on incorporating dynamic spatial datasets, real-time tourism mobility information, and continuously updated environmental observations to further enhance analytical capabilities and operational decision support.

Declarations

Funding:

"Funding not applicable".

Data Availability Statement: The data presented in this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

Author Contribution

Author Contributions: For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used "Conceptualization, B.I. and T.B.; methodology, B.I.; software, B.I.; validation, T.B. and A.R.; formal analysis, A.R.; investigation, B.I.; resources, B.I.; data curation, A.R.; writing—original draft preparation, B.I.; writing—review and editing, T.B.; visualization, A.R.; supervision, T.B.; project administration, B.I. All authors have read and agreed to the published version of the manuscript.

References

1. Hall CM, Page SJ (2014) The geography of tourism and recreation: Environment, place and space (4th ed.). Routledge. ISBN 9780415833998
2. Boers B, Cottrell S (2007) Sustainable tourism infrastructure planning. J Sustainable Tourism 15(5):492–507. https://doi.org/10.2167/jost640.0
3. Geneletti D, Dawa D (2009) Environmental impact assessment of tourism. Environ Impact Assess Rev 29(2):116–125. https://doi.org/10.1016/j.eiar.2008.06.002
4. Bahaire T, Elliott-White M (1999) The application of GIS in sustainable tourism planning. J Sustainable Tourism 7(2):159–174. https://doi.org/10.1080/09669589908667333
5. McAdam D (1999) The value and scope of geographical information systems in tourism management. J Sustainable Tourism 7(1):77–92. https://doi.org/10.1080/09669589908667326
6. Farsari I, Prastacos P (2004) GIS applications in tourism planning and management. Tourism Geographies 6(4):451–468. https://doi.org/10.1080/1461668042000280213
7. Rahman MM et al (2024) GIS-based tourism planning and spatial analysis. Sustainability 16:1452. https://doi.org/10.3390/su16041452
8. Delgado Fernández T, Delgado Fernández M (2012) SDI for tourism management. Int J Spat Data Infrastruct Res 7:148–169
9. Zhang J, Li X (2005) Design and implementation of tourism GIS. J Geog Sci 15(4):469–478. https://doi.org/10.1007/BF02873083
10. Rajabifard A, Feeney MEF, Williamson IP (2002) Future directions for SDI development. Int J Appl Earth Obs Geoinf 4(1):11–22. https://doi.org/10.1016/S0303-2434(02)00002-8
11. Nebert DD (2004) Developing Spatial Data Infrastructures: The SDI Cookbook. GSDI Association
12. Crompvoets J, Bregt A, Rajabifard A, Williamson I (2004) Assessing the worldwide developments of national spatial data clearinghouses. Int J Geogr Inf Sci 18(7):665–689. https://doi.org/10.1080/13658810410001702030
13. INSPIRE Directive (2007) Infrastructure for Spatial Information in Europe. European Commission
14. Bernard L, Kanellopoulos I, Annoni A, Smits P, Computers (2005) Environ Urban Syst, 29(1), 15–31. https://doi.org/10.1016/j.compenvurbsys.2004.05.013
15. Maguire DJ, Longley PA (2005) The emergence of geoportals and their role in spatial data infrastructures. Comput Environ Urban Syst 29(1):3–14. https://doi.org/10.1016/j.compenvurbsys.2004.05.012
16. Budhathoki NR, Nedović-Budić Z, Bruce B (2008) Reconceptualizing the role of the user of spatial data infrastructure. GeoJournal 72(3–4):149–160. https://doi.org/10.1007/s10708-008-9189-x
17. Batty M, Hudson-Smith A, Milton R, Crooks A (2010) Map mashups, Web 2.0 and the GIS revolution. Ann GIS 16(1):1–13. https://doi.org/10.1080/19475681003700831
18. Tsou MH (2011) Revisiting Web GIS: A review of directions. ISPRS Int J Geo-Information 1(1):3–18. https://doi.org/10.3390/ijgi1010003
19. Zhao P, Foerster T, Yue P (2012) The geoprocessing web. Comput Geosci 47:3–12. https://doi.org/10.1016/j.cageo.2012.04.021
20. Yao X, Li X (2013) Web-based GIS for tourism information system. J Geographic Inform Syst 5(2):163–170. https://doi.org/10.4236/jgis.2013.52016
21. Longley PA, Goodchild MF, Maguire DJ, Rhind DW (2015) Geographic Information Science and Systems, 4th edn. Wiley
22. Malczewski J (2006) GIS-based multicriteria decision analysis: A survey. Int J Geogr Inf Sci 20(7):703–726. https://doi.org/10.1080/13658810600661508
23. Eastman JR (2012) Multi-criteria evaluation and GIS. Geographical Information Systems
24. Albrecht C, Wilke T (2008) Ancient Lake Ohrid: Biodiversity and evolution. Hydrobiologia 5:181–189. https://doi.org/10.1007/S10750-008-9558-Y
25. UNESCO (2020) World heritage site: Lake Ohrid. UNESCO World Heritage Centre. https://worldheritageoutlook.iucn.org/explore-sites/natural-and-cultural-heritage-ohrid-region
26. Peterson MP (2008) International perspectives on maps and the Internet. Springer. https://doi.org/10.1007/978-3-540-72029-4
27. Izeiroski S, Idrizi B (2018) Geospatial Analysis of Suitable Locations for Exploatation of Small Hydropower and Solar Energy Potentials: A Case Study of Prespa Lake Region in Macedonia. Uluslararası Doğa Mühendislik Bilim. Derg. 2018, 12, 19–25. Available online: https://www.ijnes.org/index.php/ijnes/article/view/308
28. Gretzel U, Sigala M, Xiang Z, Koo C (2015) Smart tourism: Foundations and developments. Electron Markets 25:179–188. https://doi.org/10.1007/s12525-015-0196-8
29. Izeiroski S, Idrizi B, Lutovska M, Kabashi I (2018) GIS-based multi-criteria analysis of site suitability for exploitation of renewable energy resources. In Proceedings of the 7th International Conference on Cartography and GIS (pp. 653–665). https://www.researchgate.net/publication/326986392_GIS-BASED_MULTI_CRITERIA_ANALYSIS_OF_SITE_SUITABILITY_FOR_EXPLOATATION_OF_RENEWABLE_ENERGY_RESOURCES
30. Haklay M, Singleton A, Parker C (2008) Web mapping 2.0: The neogeography of the GeoWeb. Geogr Compass 2(6):2011–2039. https://doi.org/10.1111/j.1749-8198.2008.00167.x

31. Gutiérrez J, García-Palomares JC (2008) Distance-measure impacts on the calculation of transport service areas. Environ Plann B 35(3):480–503. https://doi.org/10.1068/b33107
32. Idrizi B, Nimani A, Pashova L (2026a) Identification of Potential Flood-Prone Areas in the Republic of Kosovo Using GIS-Based Multi-Criteria Decision-Making and the Analytical Hierarchy Process. Sustainability 18(1):359. https://doi.org/10.3390/su18010359
33. Eremchenko EN, Tikunov VS, Gorin S, Idrizi B, Radevski I (2016) Information tourism system with mixed semiotic structure. 2nd Internatonal Scientific Conference Geobalcanica 2016, 139–145. http://dx.doi.org/10.18509/GBP.2016.19
34. Xiang Z, Fesenmaier DR (2017) Big data analytics in tourism. J Travel Res 56(3):299–312. https://doi.org/10.1177/0047287516654721
35. Idrizi B, Zeqiri E, Bajrami Lubishtani F (2026b) Screening potential areas for ground-mounted solar panels by combining the Analytic Hierarchy Process (AHP) with geospatial analysis methodologies: a case study of Gjilan region, Kosova. Surv Rev J. https://doi.org/10.1080/00396265.2026.2615751
36. Haklay M (2013) Neogeography and the delusion of democratisation. Environ Plann A 45(1):55–69. https://doi.org/10.1068/a45184
37. Idrizi B, Sulejmani V, Zimeri Z Multi-scale map for three levels of spatial planning data sets for the municipality of Vitia in Kosova. In Proceedings of the 7th International Conference on Cartography and GIS, Szopol, Bulgaria (2018) 18–23 June 2018; Available online: https://www.researchgate.net/profile/Bashkim-Idrizi/publication/326986517_MULTI-SCALE_MAP_FOR_THREE_LEVELS_OF_SPATIAL_PLANNING_DATA_SETS_FOR_THE_MUNICIPALITY_OF_VITIA_IN_KOSOVA/links/5b70be8545851546c9 SCALE-MAP-FOR-THREE-LEVELS-OF-SPATIAL-PLANNING-DATA-SETS-FOR-THE-MUNICIPALITY-OF-VITIA-IN-KOSOVA.pdf?__cf_chl_tk=IMmiHLHA5UAVvjVTa1AyKigSkGPDC91cBBVRNop0mXg-1763002699-1.0.1.1-q0DOSZFpKKm_8YT8owFcJc7xir0q51sU0ZpmnG98aUE
38. Sigala M, Gretzel U (2018) Advances in social media for travel, tourism and hospitality. Routledge
39. Chen Y, Yu J, Khan S (2010) Spatial sensitivity analysis of multi-criteria weights in GIS. Environ Model Softw 25(12):1582–1591. https://doi.org/10.1016/j.envsoft.2010.06.001
40. Dissanayake D, Ranagalage M, Jayasundara J, Rajapakshe P, Herath N, Marasinghe SA, Wanninayake W, Dilanjani H, Perera A, Herath Y (2025) City Context Land 14(3):492. https://doi.org/10.3390/land14030492. A Comprehensive Bibliometric Analysis of Spatial Data Infrastructure in a Smart
41. Azizi V, Idrizi B, Nikolli P, Izeiroski S (2026) Determining of potential location for dam construction in Pograxhë region through geospatial analysis. Appl Geomatics 18:52. https://doi.org/10.1007/s12518-026-00705-2
42. Zwirowicz-Rutkowska A, Michalik A (2024) Pol Appl Sci 14(19):9117. https://doi.org/10.3390/app14199117. Spatial Data Infrastructure and Mobile Big Data for Urban Planning Based on the Example of Mikolajki Town in
43. Idrizi B, Bajrami Lubishtani F, Zeqiri E (2024) Assessment and Zoning of Areas by Risk Level of Snow Avalanches in Sharr Mountains. Tech. J. 2024, 18, 7–16. https://doi.org/10.31803/tg-20240909221359
44. Idrizi B (2006) Developing of globally homogeneous geographic data set through Global mapping project; Kartografija i geoinformacije; Zagreb. ISSN 1333-896X, https://kig.kartografija.hr/index.php/kig/article/view/546
45. Dodge M, Kitchin R (2007) Outlines of a world coming into existence. Environ Plann B 34(3):431–440. https://doi.org/10.1068/b3403ed
46. Goodchild MF (2007) Citizens as sensors: The world of volunteered geography. GeoJournal 69(4):211–221. https://doi.org/10.1007/s10708-007-9111-y

Figures

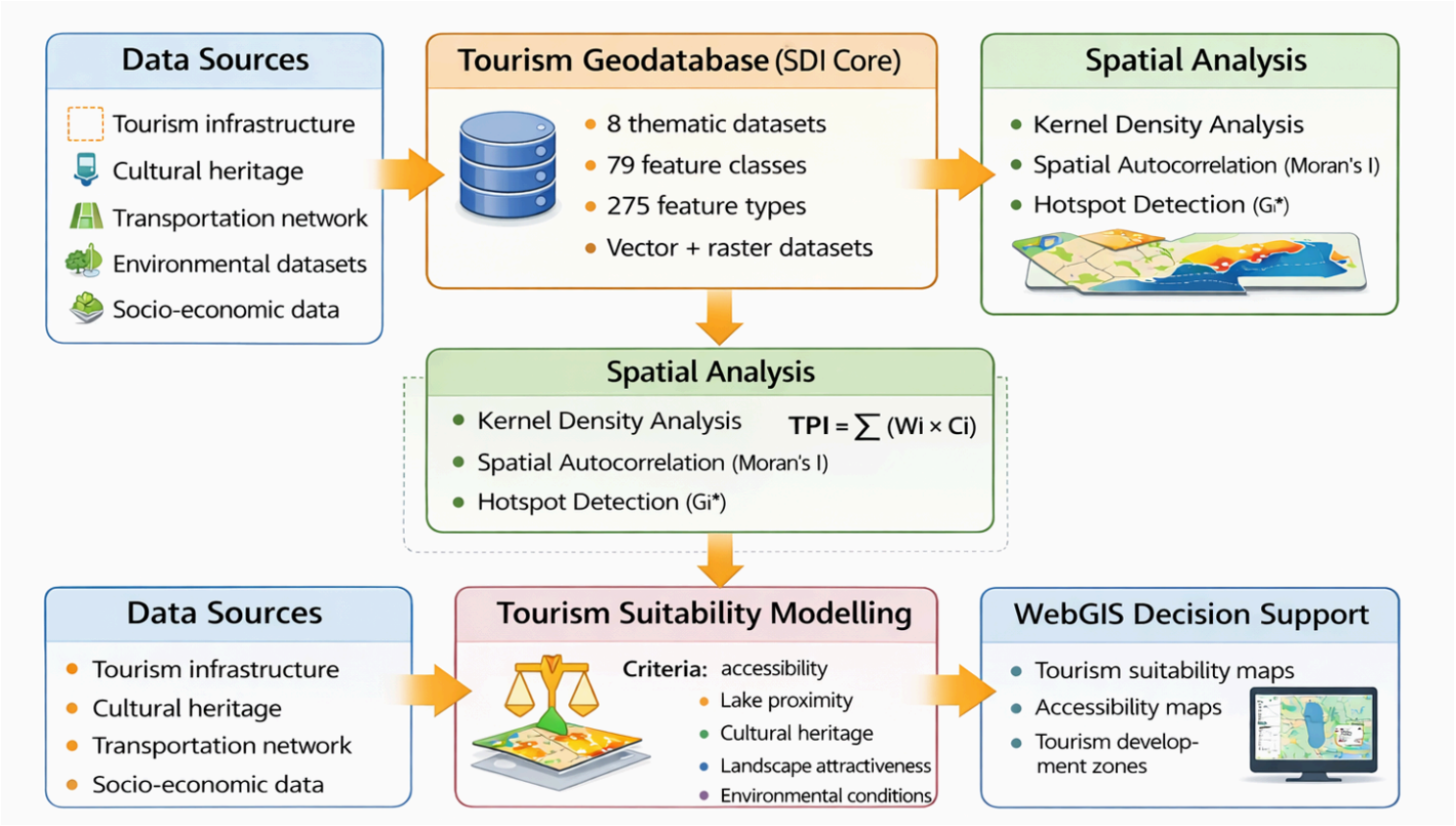

Figure 1

Analytical workflow and methodological framework of the Tourism Spatial Data Infrastructure used for spatial modelling and WebGIS-based decision support in tourism planning.

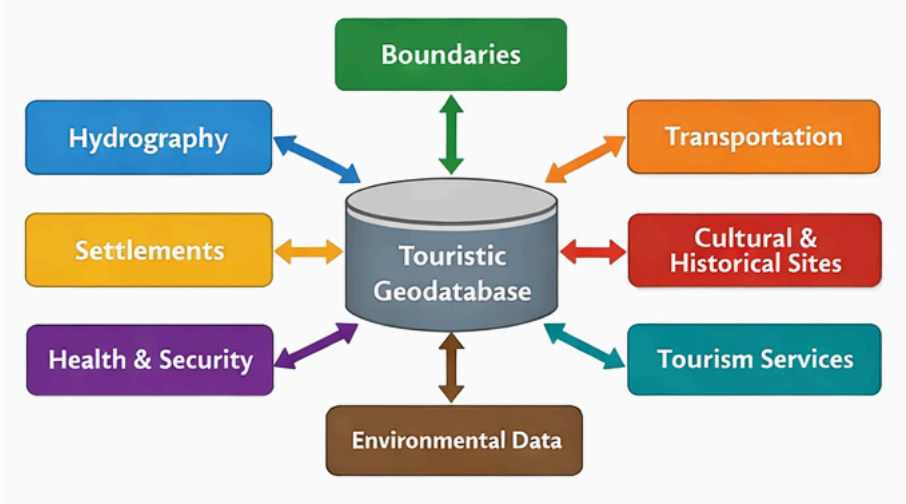

Figure 2

Conceptual internal structure of the tourism geodatabase within the proposed Spatial Data Infrastructure.

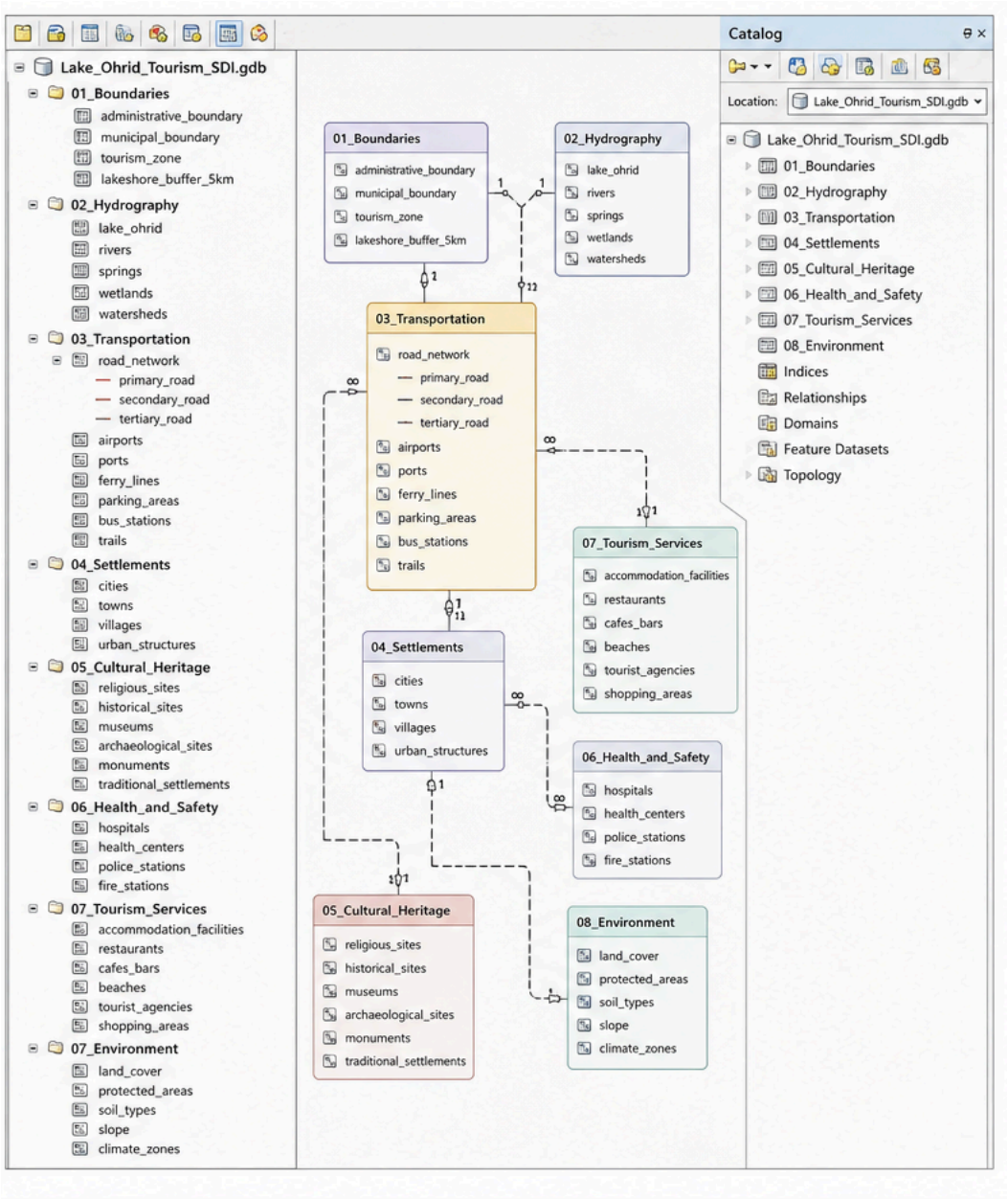

Figure 3

Implemented GIS geodatabase structure of the Lake Ohrid Tourism Spatial Data Infrastructure showing thematic feature datasets and their logical organization.

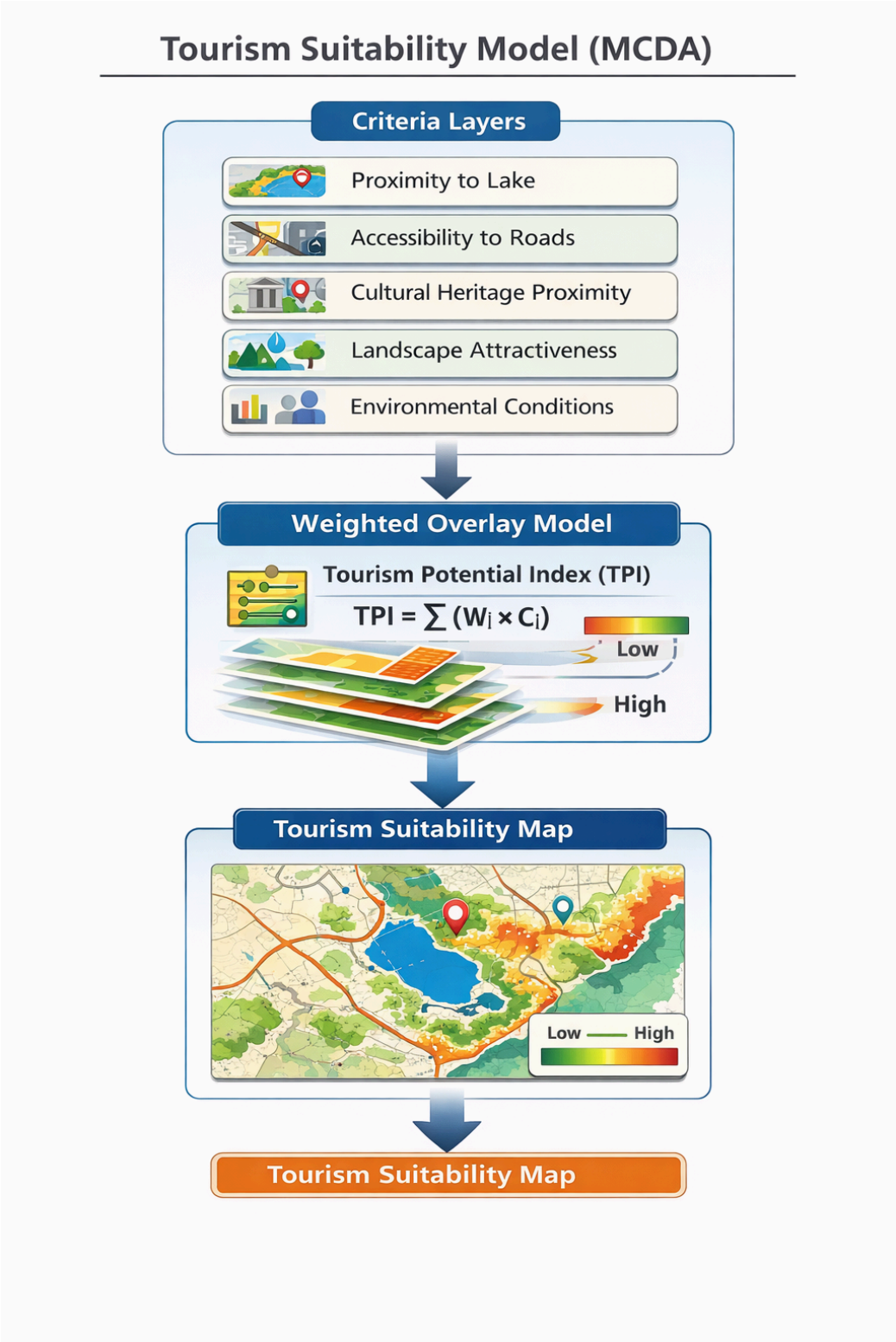

Figure 4

Conceptual workflow of the multi-criteria tourism suitability modelling approach

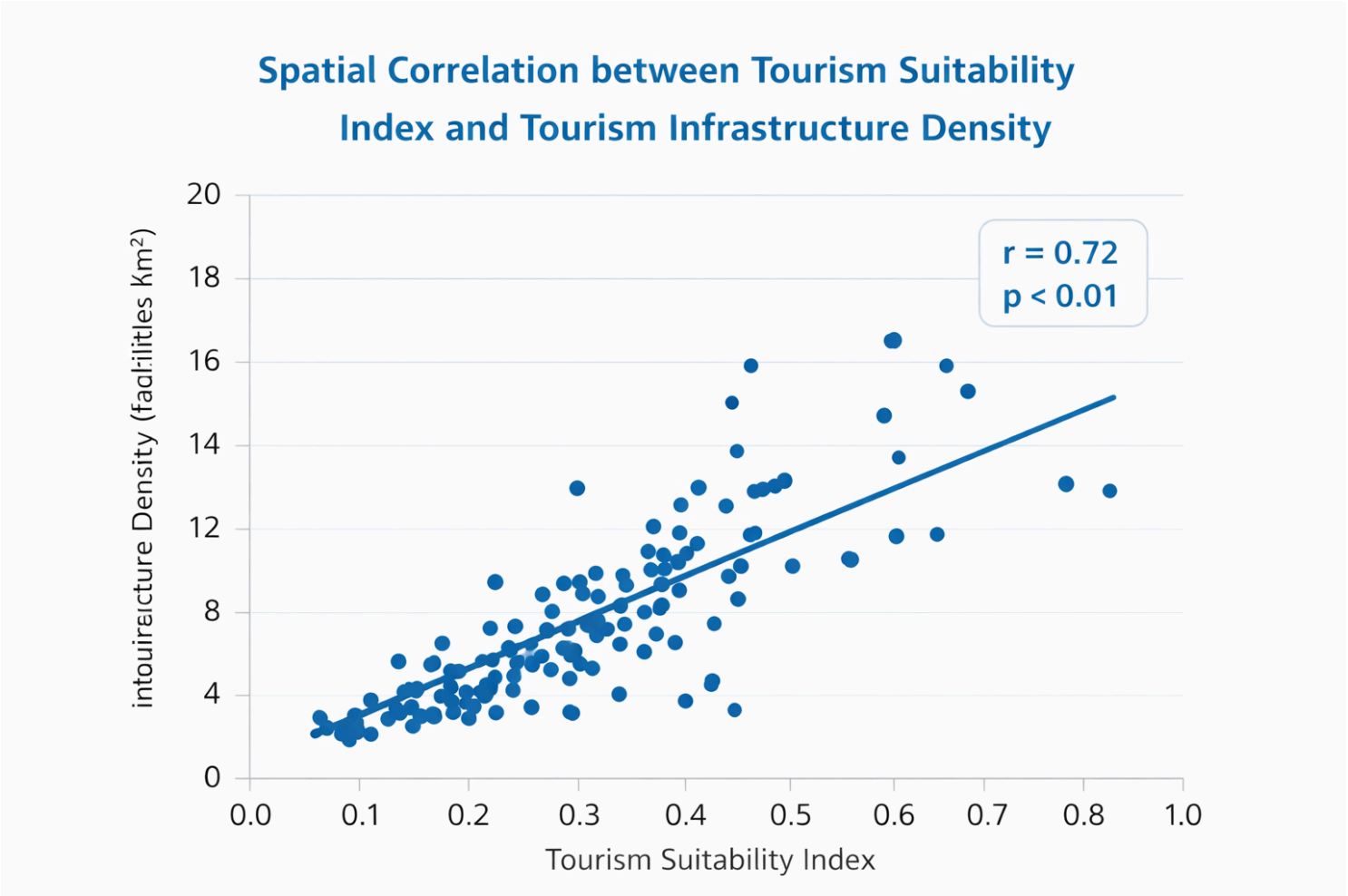

Figure 5

Spatial correlation between tourism suitability index and tourism infrastructure density

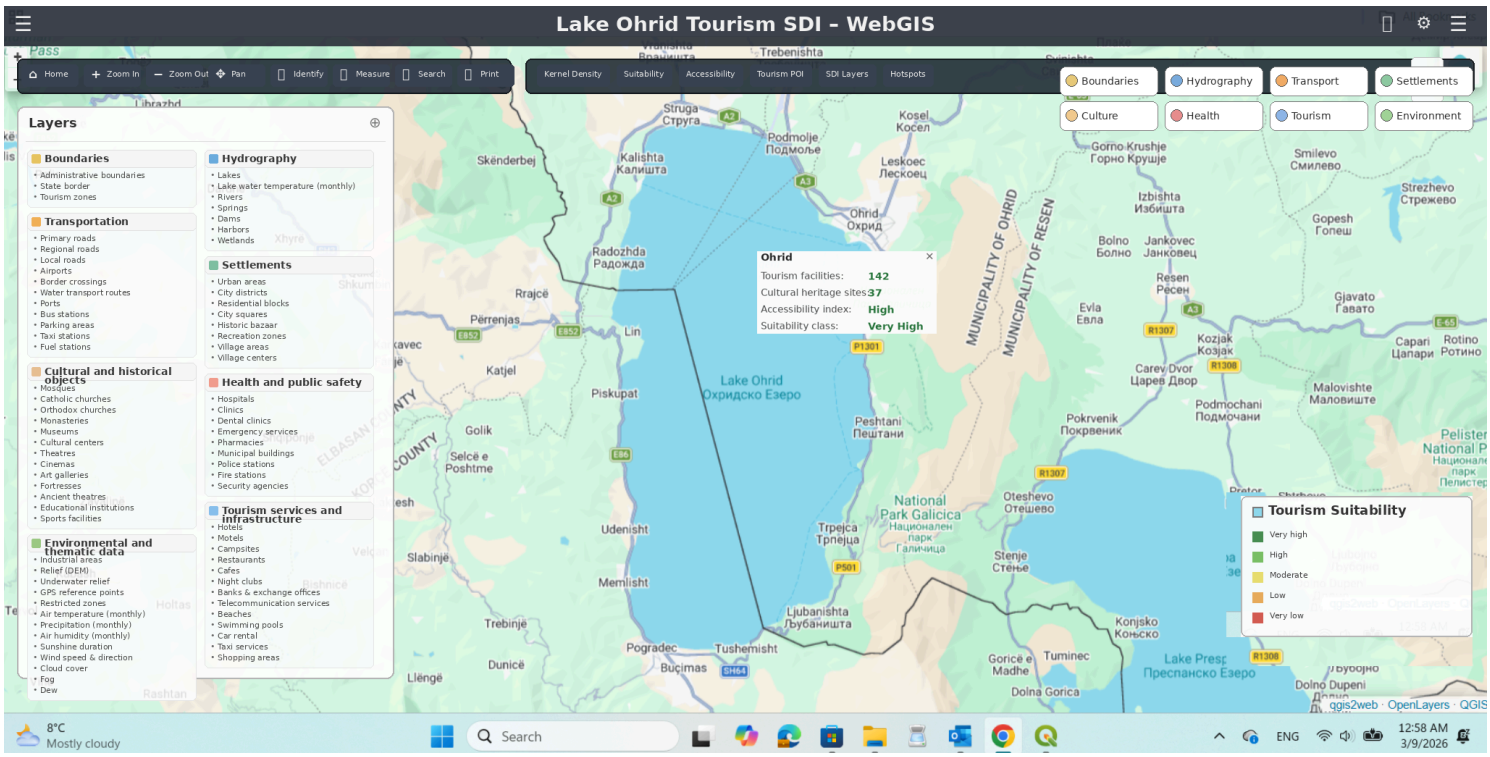

Figure 6

Operational WebGIS environment of the Lake Ohrid Tourism Spatial Data Infrastructure integrating thematic datasets, tourism suitability modelling results, and interactive visualization and querying functions for spatial decision support.