Italian Journal of Engineering Geology and Environment
www.ijege.uniroma1.it

GEOSTATISTICAL ANALYSIS OF THE SPATIAL DISTRIBUTION OF ENVIRONMENTAL DATA: A SURVEY ON METHODS AND APPLICATIONS

Sander KOVACI(*) & Alfred LAKO(**)

(*)Polytechnic University of Tirana, Department of Mathematical Engineering -Tirana (Albania)
(**)Polytechnic University of Tirana - Department of Environmental Engineering - Tirana (Albania)
Corresponding author: lakoalfred025@gmail.com

EXTENDED ABSTRACT

L'analisi della distribuzione spaziale dei dati ambientali è diventata sempre più critica nelle moderne scienze naturali e tecniche, consentendo ai ricercatori di comprendere modelli e meccanismi di interazione tra i processi naturali. La ricerca ha seguito le linee guida PRISMA per la revisione sistematica della letteratura, esaminando inizialmente articoli identificati da database Web of Science, Scopus e Google Scholar, selezionando infine studi che soddisfacevano rigorosi criteri di inclusione in termini di rigore metodologico, pertinenza ai metodi geostatistici e integrazione con le moderne tecnologie geospaziali. La geostatistica è un campo interdisciplinare che integra statistica, geografia e tecnologie dell'informazione per analizzare i dati spaziali, basandosi sul concetto di dipendenza spaziale articolato nella Legge di Tobler. Lo studio ha esaminato concetti geostatistici fondamentali, tra cui valore medio, varianza, autocorrelazione spaziale misurata attraverso l'indice I di Moran e LISA, che collettivamente consentono una valutazione completa delle caratteristiche globali e locali della distribuzione dei dati nello spazio. La ricerca ha esaminato diversi metodi di interpolazione, con il kriging che si è rivelato l'approccio più sofisticato basato sull'analisi del variogramma, offrendo una precisione superiore rispetto a metodi più semplici come la ponderazione della distanza inversa (IDW) e l'interpolazione spline. Inoltre, lo studio ha esplorato modelli di regressione, tra cui regressione lineare, regressione multipla e regressione ponderata geograficamente (GWR), insieme a modelli di classificazione discreta come la regressione logistica e gli alberi decisionali, dimostrandone l'efficacia nella modellazione di relazioni spaziali complesse e fenomeni categoriali. L'integrazione del GIS con i metodi geostatistici rappresenta un progresso cruciale nelle capacità di analisi spaziale. La ricerca ha valutato sia piattaforme open source, come QGIS, R (con librerie che includono sp, sf, gstat e geoR), sia software commerciali, tra cui ArcGIS con il suo modulo Geostatistical Analyst e MATLAB, dimostrando come questi strumenti facilitino l'elaborazione, la visualizzazione e la modellazione complete dei dati. Python si è affermato come un linguaggio di programmazione particolarmente versatile per l'analisi spaziale, con librerie come GeoPandas, PySAL, Rasterio e Shapely che offrono ampie funzionalità per la gestione di dati sia vettoriali che raster. Piattaforme basate su cloud, tra cui Google Earth Engine (GEE), ArcGIS Online, AWS Location Service, QGIS Cloud e Mapbox, sono state analizzate per la loro capacità di elaborare immagini satellitari su larga scala, integrare dati multi-sorgente e creare mappe interattive ad alta risoluzione necessarie per un monitoraggio ambientale accurato. L'applicazione pratica dei metodi geostatistici è stata dimostrata in tre ambiti ambientali critici: inquinamento atmosferico, valutazione della qualità del suolo e monitoraggio delle risorse idriche. Per l'analisi della qualità dell'aria, l'interpolazione spaziale ha rivelato che le concentrazioni di biossido di azoto nelle aree ad alto traffico hanno raggiunto i 45-60 μg/m³, superando le soglie dell'OMS, mentre i livelli di PM2.5 nelle zone industriali hanno raggiunto i 30-42 μg/m³. Il metodo kriging ha raggiunto un'elevata accuratezza predittiva con valori RMSE compresi tra 2.5 e 4.2, superando significativamente l'interpolazione IDW (RMSE=5.8-7.1). L'analisi della contaminazione del suolo ha identificato concentrazioni di piombo pari a 140-180 mg/kg nelle aree industriali e livelli di cadmio pari a 1.2-2.5 mg/kg nelle regioni agricole con uso intensivo di fertilizzanti, entrambi superiori agli standard normativi. La valutazione della qualità dell'acqua ha rivelato concentrazioni di nitrati pari a 40-65 mg/L nei fiumi in prossimità delle zone agricole, superando le concentrazioni massime consentite, mentre la contaminazione da metalli pesanti da fonti industriali ha mostrato livelli elevati di piombo, cadmio e mercurio. L'analisi di autocorrelazione spaziale ha confermato modelli di inquinamento a cluster con valori di I di Moran compresi tra 0.64 e 0.71, indicando fonti di inquinamento stabili e facilitando strategie di intervento mirate.

Il presente studio ha identificato diversi progressi critici nella metodologia geostatistica tuttavia, la combinazione di approcci geostatistici tradizionali con metodi computazionali contemporanei migliora l'accuratezza delle previsioni, consente il monitoraggio in tempo reale e facilita l'elaborazione automatizzata dei dati per valutazioni ambientali su larga scala. I modelli di classificazione hanno raggiunto un'accuratezza sostanziale, con una regressione logistica che ha raggiunto l'84% e alberi decisionali del 79% nell'identificazione delle aree contaminate in tre categorie di inquinamento: zone criticamente inquinate, aree moderatamente inquinate e territori ecologicamente stabili.

Nonostante i significativi progressi metodologici, sono stati riscontrati alcuni limiti,quali l'orientamento teorico che richiede una validazione empirica in diversi contesti ambientali e la necessità di una maggiore integrazione con tecnologie emergenti come l'intelligenza artificiale. Gli sviluppi futuri includono l'adattamento dei metodi geostatistici a specifiche condizioni ecologiche, lo creazione di algoritmi che tengano conto delle caratteristiche degli ecosistemi locali e la creazione di nuovi approcci per combinare l'analisi spaziale con le previsioni ambientali.

IJEGE, 2 (2025): 75-92, DOI: 10.4408/IJEGE.2025-02.O-05
E-ISSN 2035-5688 | ISSN 1825-6635 / ©Author(s)

ABSTRACT

The aim of this study was to examine the methods, tools, and platforms used for the analysis of spatial data, as well as to assess their potential for solving environmental challenges. The study considers geostatistical methods, including kriging, variogram analysis, semivariance, the Thiessen polygon method, inverse distance weighting (IDW) interpolation, and regression models such as linear regression, multiple regression, and geographically weighted regression (GWR). A literature search turned up 52 peer-reviewed and indexed articles on methods including regression models, variogram analysis, and kriging. The selection criteria included: (1) relevance to geostatistical analysis of environmental data, (2) methodological rigor, (3) publication in high-impact peer-reviewed journals, and (4) citation frequency indicating scientific significance. These studies highlight the effectiveness of geostatistical methods, geospatial platforms, and Python in environmental monitoring and predictive modeling. For classification tasks, logistic regression and decision trees were examined. The study results demonstrate that the application of modern geostatistical methods allows for the identification of spatial distribution patterns of environmental data and improves prediction accuracy. In particular, it was found that the spatial autocorrelation index effectively determines areas with high levels of similarity in environmental parameters, while local indicators of spatial association (LISA) help identify regional clusters with high pollution intensity or other anomalous characteristics. It was demonstrated that the use of spatial modelling platforms, such as Geographic Information System (GIS) software like ArcGIS and Quantum GIS (QGIS), along with the Python programming language and spatial data analysis libraries such as GeoPandas and the Python Spatial Analysis Library (PySAL), significantly enhances the effectiveness of environmental phenomenon analysis. The integration of satellite image data with geostatistical methods was found to contribute to the creation of more accurate maps for forecasting environmental risks. The proposed approaches demonstrate significant potential for environmental monitoring and natural resource management, enhancing the understanding of spatial patterns and serving as a basis for further research in this field.

Keywords: *geostatistics, geostatistical analysis, variogram, semivariogram, kriging methods, geographic information system*

INTRODUCTION

The analysis of the spatial distribution of environmental data is a key task in modern natural and technical sciences, as it enables the understanding of patterns and mechanisms of interaction between natural processes. The growing number of environmental challenges, such as climate change, air, soil, and water pollution, increases the need for effective methods of analysis and forecasting. Simultaneously, the increased availability of spatial data through satellite technologies and Geographic Information System (GIS) creates new opportunities for research (Ramoul *et alii*, 2022).

Geostatistical analysis has become an essential tool for assessing the spatial distribution of environmental data and studying the impact of various factors on the environment. Contemporary research has shown that the use of geostatistical methods not only facilitates effective spatial process modelling but also provides accurate estimates for decision-making. Scientific studies confirm the universality and effectiveness of GIS and geostatistical methods across various fields of spatial data analysis. For example, Koto *et alii* (2022) utilised GIS to analyse water quality in the Karavasta Lagoon (Albania), focusing on determining concentrations of heavy metals and nutrients. The study contributed to assessing environmental risks and identifying pollution zones, enabling recommendations to minimise anthropogenic impact. In research by Shehu (2023), the spatial distribution of macro- and microelements in the soils of the Vushtrri region (Kosovo) was analysed. Using geostatistical methods, the author identified zones of nutrient deficiency and surplus, which is crucial for sustainable agricultural development and land use optimisation. Meanwhile, Berila & Isufi (2021) applied GIS to analyse vertical terrain fragmentation in the Drenica River basin (Kosovo). Their study highlighted the importance of spatial analysis in examining geomorphological characteristics and assessing the impact of terrain on hydrological processes. Finally, the work of Mema *et alii* (2024) focused on modelling the spatial variability of criteria affecting land suitability for agricultural use. Using GIS, the authors created models to determine optimal areas for crop cultivation, considering climatic and agronomic conditions.

Previous studies have examined various aspects of applying geostatistical methods to analyse the spatial distribution of environmental parameters. The dynamics of heavy metal distribution in the environment were studied using spatial mapping methods, variability analysis, and pollution source identification (Sylva *et alii*, 2024). It was found that these approaches effectively assess pollution sources and model the spatial variability of contaminants. The general application of geostatistics to environmental tasks has been explored, with its potential for predicting environmental changes using mathematical models adapted to different scenarios analysed (Zanini & D'Oria, 2024). Special attention was given to developing models that consider complex interrelationships within natural systems. Research has been conducted on predicting the spatial distribution of heavy metals in soils using limited field measurement data and multi-source information. New interpolation approaches have been developed to provide more accurate predictions even under data constraints (Sun *et alii*, 2024).

The application of hybrid machine learning and geostatistics methods for spatial forecasting of invasive plant spread has proven effective in identifying relationships between environmental factors and species distribution (Shen *et alii*, 2024; Phoophathong *et alii*, 2025). This research has highlighted the potential of such approaches for predicting environmental risks. An assessment of groundwater quality in rural areas was conducted using geostatistical methods and WebGIS, enabling the integration of spatial analysis with modern online visualisation technologies (Balla *et alii*, 2024). The results indicate the practicality of these methods for monitoring and managing environmental resources. Within the framework of soil modelling, geostatistical methods and remote sensing were integrated, allowing for the creation of accurate predictive maps for precision agriculture (Hilal *et alii*, 2024). This approach proved effective in accounting for soil variability at the regional level. Research into the application of geostatistical methods for evaluating reservoir characteristics demonstrated their effectiveness when combined with petrophysical analysis for resource optimisation in the oil industry (Esiri *et alii*, 2024).

The development and application of GIS for spatial analysis in environmental geochemistry were examined. The analysis showed that these approaches are indispensable in the context of big data for studying the dynamics of natural processes and pollution (Xu & Zhang, 2023). Additionally, the spatial distribution of groundwater quality parameters was studied using geostatistical models, revealing regional characteristics and gaps in monitoring (Farzaneh *et alii*, 2022). Spatial analysis of soil properties in agricultural regions was carried out using GIS methods, which demonstrated their effectiveness in accounting for spatial variability to optimise agricultural practices (Khan *et alii*, 2021).

Despite significant progress in the development of geostatistical methods, gaps remained in the scientific literature, particularly concerning their application in specific environmental conditions. The effectiveness of various methods at different scales of analysis and their integration with modern automated data processing tools had not been sufficiently studied. For example, spatial interpolation methods such as kriging, although widely used, required further refinement for working with complex ecosystems where non-linear relationships and multiple influencing factors exist. The aim of this study was to analyse contemporary geostatistical methods and platforms in terms of their effectiveness in solving environmental challenges, as well as to assess their adaptability to complex ecological systems and integration with automated data analysis tools.

MATERIALS AND METHODS

A systematic approach was used to analyse modern methods, tools, and platforms for geostatistical analysis of the spatial distribution of environmental data. To ensure a comprehensive and transparent literature search, the study followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The literature search was conducted in three widely recognised academic databases: Web of Science, Scopus, and Google Scholar, to ensure a comprehensive selection of peer-reviewed articles. These databases were chosen for their broad coverage of scientific journals and their ability to access high-quality research publications. Among the terms used in the search were "geostatistics", "spatial data analysis", "kriging", "GIS", "environmental data", "pollution mapping", and "spatial autocorrelation". These keywords were chosen in order to encompass a broad spectrum of pertinent research pertaining to data analysis methodologies, environmental monitoring, and geostatistical approaches. To obtain thorough results and narrow search queries, boolean operators (AND, OR) were used.

In order to identify the most influential studies, the following inclusion criteria were used to select articles: (1) relevance to geostatistical methods or spatial data analysis in environmental studies; (2) methodological rigour and clarity in reporting; (3) publication in peer-reviewed, high-impact journals; and (4) citation frequency. 51 articles were chosen for in-depth study after duplicates were eliminated using these criteria. The final choice was made after a thorough assessment of entire texts, abstracts, and titles to make sure they matched the goals and parameters of the study.

A total of 254 articles were initially identified through the literature searche (Fig. 1). After removing 62 duplicates, 192 articles remained for further screening. During the title and abstract screening phase, 34 studies were excluded due to irrelevance, either not addressing geostatistical methods, GIS, or spatial data analysis, or focusing on unrelated topics. In the full-text screening phase, 106 studies were excluded for failing

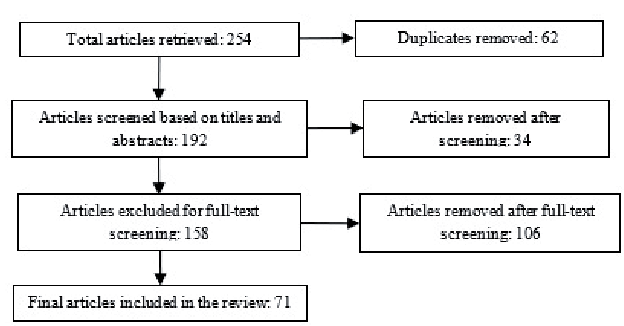

Fig. 1 - Flowchart of literature search and selection process

to meet the inclusion criteria. These studies were excluded for being irrelevant to geostatistical methods and environmental data analysis, lacking integration with modern geospatial technologies (such as GIS or remote sensing), being written in a language other than English, or having methodological issues such as the absence of practical applications or detailed descriptions of methods. Ultimately, 52 studies were included in the final review, meeting the criteria of being relevant to geostatistical analysis using GIS and programming tools, and incorporating modern platforms like ArcGIS, Quantum GIS (QGIS), and Python for environmental data analysis.

The study also described spline interpolation methods, including cubic splines, which enable the creation of smooth models for spatial data based on existing measurement points. Among the tools used for analysis, the most notable were modern software platforms: ArcGIS (Esri, 2025), QGIS (Quantum Geographic Information System (QGIS) Development Team, 2025) and MATLAB (MathWorks, 2025). Other important tools were GeoDa (2025) and Surfer (Golden Software, 2025). The study also utilised R programming languages (R Core Team, 2025) with the libraries sp, sf, gstat, and geoR. In addition, Python was applied using specialised geospatial libraries such as GeoPandas (2025), Python Spatial Analysis Library (PySAL) (2025), Shapely (2025), and Matplotlib (2025).

The use of cloud services such as Google Earth Engine (GEE) (2025), Amazon Web Services (AWS) Location Service (2025), Quantum Geographic Information System (QGIS), Cloud (2025), and Mapbox (2025) enabled data integration, satellite image processing, and the creation of high-resolution maps necessary for accurate monitoring of pollution and environmental changes.

The study examined various geostatistical analysis methods for assessing the level of pollution in atmospheric, soil, and water resources. The kriging method was used for interpolating pollutants in missing data points, allowing for accurate predictions of spatial pollution distributions. Semivariograms and variogram analysis were applied to identify the spatial variability of pollutants and assess their dependence on the distance between points. These methods were mainly used for analysing soil and water data, where spatial correlation is a key aspect.

The inverse distance weighting (IDW) interpolation method was used for modelling pollution in areas with a limited dataset, which was particularly relevant for water resources. This approach is based on Tobler's law, enabling the creation of models that accurately reflect real pollution conditions in spatial proximity.

Thiessen polygons were employed for classifying territories by pollution levels, particularly atmospheric and soil pollution, where it was necessary to identify zones with high pollution impact. Regression models, including linear regression, multiple regression, and geographically weighted regression (GWR), were used to model the relationships between pollution and other factors such as transport, industrial emissions, and weather conditions, allowing for the assessment of local effects of these factors.

For pollution classification tasks, logistic regression and decision tree methods were applied. These were used to classify territories by pollution levels, particularly soil and water pollution, to identify complex spatial patterns.

Spline interpolation, particularly cubic splines, was also applied for interpolating environmental parameters such as soil moisture or pollutant concentrations in water, enabling the creation of smooth spatial models for locations with limited data availability.

The selection of methods was determined by their ability to address specific tasks. For instance, the kriging method was chosen for its capacity to account for spatial dependence and ensure high prediction accuracy. IDW was suitable for analysis in data-limited conditions, whereas regression methods allowed for the assessment of external factors' influence on pollution. The Thiessen polygon method was selected for delineating clear pollution zones, which is crucial for management decision-making. Spline interpolation ensured model smoothness, which is particularly relevant for the analysis of continuous environmental parameters.

All these methods were applied within the framework of software platforms such as ArcGIS, QGIS, MATLAB, GeoDa, and Surfer, as well as using programming languages R and Python, enabling the execution of the necessary spatial analysis and modelling. The evaluation of the selected methods' potential was based on the following criteria: prediction accuracy (assessed using root mean square error (RMSE) and the coefficient of determination (R^2)), spatial resolution (determined by the number of control points and pixel size in the case of raster data), computational efficiency (measured by data processing time for each method), and adaptability to different data types (evaluated by comparing results for various environmental indicators, such as soil quality, atmospheric air, and water resources). Cloud services such as ArcGIS Online, GEE, AWS Location Service, QGIS Cloud, and Mapbox were used for data integration, satellite image processing, and the creation of high-detail maps, facilitating accurate environmental monitoring and forecasting.

RESULTS AND DISCUSSION

Geostatistics: an interdisciplinary data analysis tool

Geostatistics is an interdisciplinary field that integrates statistics, geography, and information technology to analyse spatial data. It studies the patterns of spatial variation in natural and anthropogenic phenomena, providing tools for their modelling and prediction. The foundation of geostatistics is the concept of spatial dependence, which suggests that values of a variable at closely located points tend to be more similar than those at distant points. This phenomenon, known as Tobler's Law - "everything is related to everything else, but near things are more related than distant things" - is fundamental to all geostatistical methods. Spatial dependence not only allows for the analysis of existing data but also facilitates predictions for areas where data have not been collected.

One of the key tools in geostatistics is the variogram - a graphical representation that reflects the relationship between the variance of a variable's values and the distance between points (Fig. 2). The variogram helps identify the spatial structure of data and determine the main parameters of this dependence: range, sill, and nugget. The range defines the maximum distance at which spatial correlation is still observed. The sill represents the stable variance value reached at large distances. The nugget indicates variation caused by local factors or measurement errors. These parameters are critically important for model construction and the selection of an appropriate analysis method (Webster & Oliver, 2007).

phenomenon where spatial dependencies vary depending on direction. In environmental data, this can be linked to the presence of rivers, mountain ranges, or other linear natural structures. Accounting for anisotropy in modelling allows for more accurate results that are better adapted to real geographical conditions.

Geostatistical principles encompass several key stages of analysis. The first step involves defining the spatial structure of data, including the analysis of variation characteristics, variogram construction, and selecting an appropriate dependency model. The next stage is interpolation, which allows for estimating variable values at locations where data are absent. This step is particularly crucial for ecological applications, land use, and natural resource monitoring. Additionally, visualising results is a fundamental part of geostatistics. The creation of spatial distribution maps enables the clear representation of patterns identified during analysis and facilitates decision-making in natural resource management.

Geostatistics serves as an essential tool for analysing complex natural and social processes. Its methods find applications in ecology, geology, agriculture, urban planning, as well as in climate change forecasting and natural resource management. By accounting for

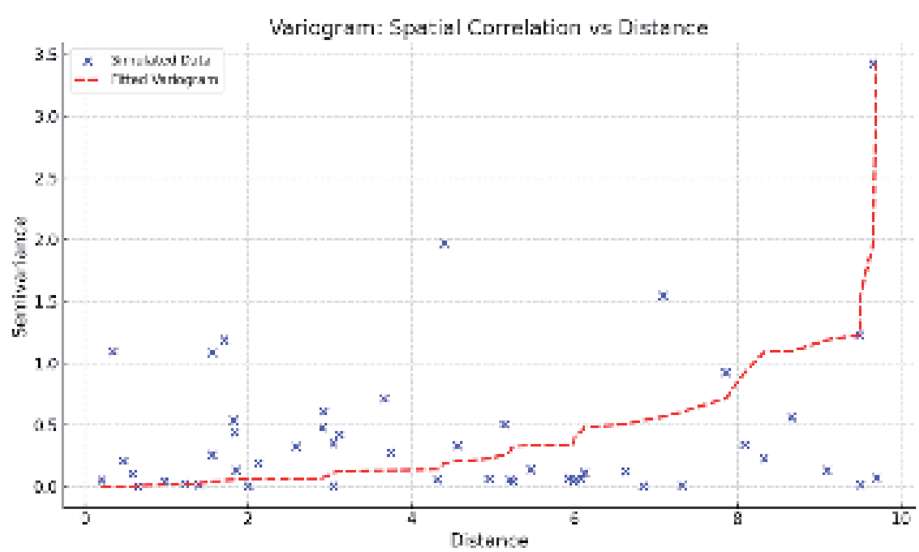

Fig. 2 - An example of variogram demonstrating how spatial correlation varies with distance

Geostatistical methods encompass various approaches to data interpolation, with kriging being the most widely used. This method enables the prediction of variable values at unsampled locations while considering spatial dependence, minimising prediction errors. Unlike simpler methods such as IDW or splines, kriging is based on the variogram, making it more precise and adaptive to the data's specifics (Cressie, 1993).

An important aspect of geostatistics is anisotropy - a

spatial dependencies and its flexibility across different scales, geostatistics provides opportunities for more accurate data analysis and well-grounded decisions across numerous disciplines. The analysis of spatial data distribution is one of geostatistics' key tasks, allowing for the investigation of how specific characteristics vary within the studied territory. Various indicators and metrics are used to assess both overall distribution characteristics and local spatial relationships (Chilès & Delfiner, 2012).

Among the fundamental concepts are mean value and variance. The mean value is used to determine the central tendency of a spatial phenomenon, such as the average pollution level in a region. Variance helps assess how much variable values deviate from the mean, which is essential for understanding spatial variability. However, these global metrics do not account for spatial dependencies and are usually complemented by other indicators. Spatial autocorrelation is a crucial tool for evaluating relationships between variable values at different locations (Fig. 3). Moran's I index is a widely used measure that determines the nature of distribution: clustered, random, or uniform. For example, in studies on vegetation cover distribution, this index helps identify areas with a high concentration of similar values.

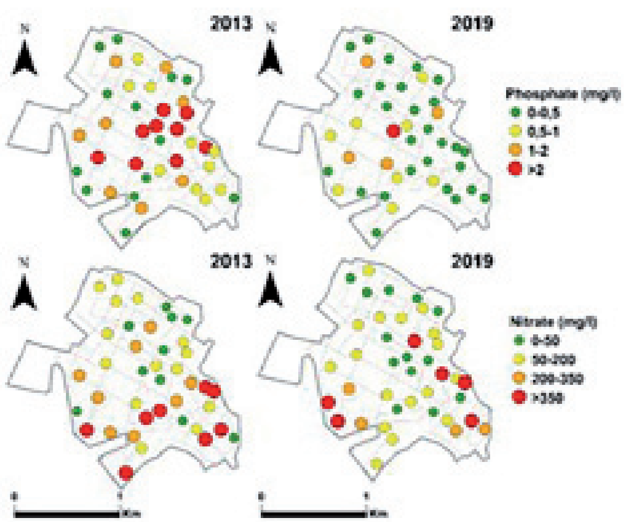

Fig. 3 - Example of spatial distribution of phosphate and nitrate values (Source: Balla et alii (2020))

Local indicators of spatial association (LISA) expand the analysis by enabling the identification of spatial patterns at the level of individual regions. These are particularly valuable for detecting local clusters, such as disease hotspots, areas of intensive land use, or regions with high population density. Another critical tool is the analysis of variations through semivariance. This measure is used for constructing variograms that demonstrate how variance between variable values depends on the distance between observation points. Variograms help determine the range of spatial dependence, which is crucial for modelling environmental processes (Cressie, 1993). Furthermore, the specificity of spatial data often requires consideration of anisotropy - the dependence of spatial relationships on direction. For example, in atmospheric studies, accounting for wind direction is crucial for the correct analysis of pollution dispersion. The core concepts and principles of geostatistics are summarised in Table 1.

These indicators and metrics are fundamental tools of geostatistical analysis, enabling the investigation of both global and local characteristics of data distribution in space. They are essential for assessing ecological, social, and economic phenomena, ensuring more precise and scientifically grounded decision-making.

Characteristics and specificity of spatial data

Spatial data possess a number of features that significantly affect the process of their analysis and the interpretation of results. One of the key characteristics of spatial data is their spatial correlation. Spatial correlation implies that the values of variables at points that are located close to one another tend to be similar. This property is based on Tobler's principle, which states, "everything is related to everything else, but closer objects are more related." However, this very dependency creates challenges for the application of traditional statistical methods, which are based on the assumption of the independence of observations.

Scaling is another important feature of spatial data. The

Concept/indicator	Description	Example of application
Mean Value (Koto et al., 2022)	Determines the average value of a variable within the studied area.	The average level of air pollution in the region.
Variance (Sylva et al., 2024)	Indicates the variability of a variable's values in space.	Assessment of the dispersion of heavy metal concentrations in soil.
Moran's I Index (John et al., 2021)	A spatial autocorrelation measure that characterises the relationship between values depending on their location.	Identification of clusters of urban areas with high or low population density.
Semivariance (Cressie, 1993)	Assesses the dependence between variable values based on the distance between points.	Construction of variograms for modelling soil moisture distribution.
Anisotropy (Webster and Oliver, 2007)	Spatial dependence that varies depending on direction.	Analysis of the direction of pollutant dispersion in a river system.
LISA (Singh and Sarma, 2023)	LISA for detecting local clusters or anomalies.	Identification of hotspots of specific plant species distribution in protected areas.

Tab. 1 - Key concepts and principles of geostatistics (Source: compiled by the authors based on Cressie (1993), Koto et alii (2022), Sylva et alii (2024), John et alii (2021), Webster & Oliver (2007), Singh & Sarma (2023)) choice of the scale at which the analysis is conducted largely determines the nature of the patterns detected. For instance, at the local level, small-scale factors such as local landscape features may dominate, while at the global level, large-scale processes such as climate change become more apparent. The analysis scale influences spatial dependencies: phenomena that show strong correlation at one scale may be negligible or entirely absent at another.

The geographical specificity of spatial data is particularly significant. The geographical environment in which the data are located directly affects their distribution and dependencies. For example, topography, climatic conditions, soil characteristics, or anthropogenic factors can significantly modify the patterns observed in different regions. The same phenomena in different geographical conditions may have completely different explanations. Thus, spatial analysis requires the consideration of the context of the territory being studied. To systematise these aspects, Table 2 is proposed.

Characteristic	Description	Example of impact on analysis
Spatial Dependence (Minh et al., 2024)	The relationship between variable values at closely located points.	Analysis of pollution dispersion, where concentration values in neighbouring areas exhibit strong similarity.
Scaling (Scale) (Cressie, 1993)	The impact of analysis scale on identified patterns.	Local processes (microclimate) and global trends (climate change) may have different correlations.
Geographical Specificity (Patra et al., 2025)	The influence of natural and anthropogenic factors on data distribution patterns.	The distribution of forests in mountainous regions depends on altitude and climatic zones.

Tab. 2 - Key concepts and principles of geostatistics (Source: compiled by the authors based on Cressie (1993), Minh et alii (2024), Patra et alii (2025))

Considering these aspects is crucial for the effective application of geostatistical methods, as they ensure the correct interpretation of data and their adaptation to specific spatial conditions.

Geostatistical analysis methods allow for a detailed study of spatial patterns and the creation of predictions necessary for decision-making in environmental protection, agriculture, and other fields (Khoso *et alii*, 2025). Each method has its unique methodology, making it suitable for specific tasks depending on the characteristics of the data and the research objectives. Kriging is a complex yet powerful tool based on variograms, which describe the spatial correlation between observation points (Soto *et alii*, 2022). It minimises forecasting errors, making it indispensable for tasks with high accuracy requirements. Splines facilitate the creation of continuous models with smooth transitions, which is particularly useful for phenomena with high variability, such as topographical or climatic data. They help avoid abrupt changes between points while preserving the integrity of the model.

The Thiessen polygon method, despite its simplicity, provides an intuitive representation of the spatial distribution of the influence of each point. It is used for creating models of resource accessibility or assessing the impact zone of objects. IDW is a more intuitive approach that assigns weight based on the distance between points. This method is suitable for phenomena with smooth variation, such as pollution levels or temperature. However, its limitations become apparent when dealing with complex dependencies (Table 3).

Method	Main principle	Application area
Kriging (Du et al., 2021)	Uses spatial correlation to predict values while minimising errors.	Resource estimation, air or soil pollution forecasting.
Splines (Cressie, 1993)	Constructs smooth curves or surfaces for data interpolation.	Topography, climate parameter modelling.
Thiessen Polygons (Thiesen and Ehret, 2022)	Divides space into zones of influence based on the nearest observation points.	Resource accessibility analysis, infrastructure placement.
IDW (Raco et al., 2021)	Uses weighted averages of neighbouring points, with weights inversely proportional to distance.	Distribution of continuous phenomena such as pollution levels or precipitation.

Tab. 3 - Comparison of key geostatistical analysis methods (Source: compiled by the authors based on Cressie (1993), Du et alii (2021), Raco et alii (2021), Thiesen & Ehret (2022))

Geostatistical analysis methods, such as kriging and splines, provide highly accurate predictions and modeling of complex spatial phenomena due to their ability to account for spatial correlation between values. These methods are used to create precise models that can predict values at unsampled locations, which is important for forecasting various ecological, geographical, or economic parameters. However, for simpler tasks involving the analysis of spatial patterns, methods such as Thiessen polygons and IDW interpolation are applied. These methods are less complex to implement but are effective for basic data analysis where high accuracy or complex dependencies between points are not required.

Since geostatistical methods focus on identifying spatial dependencies, regression and discrete models play an important role, as they not only allow for the analysis of relationships between variables but also enable the modeling of processes underlying spatial phenomena. For example, linear regression is used to model simple relationships between variables, while multiple regression allows for the consideration of several factors simultaneously, enabling more accurate predictions of variable behaviour based on their interrelationships. To analyse spatial phenomena with consideration of geographic specifics, GWR models are employed, which account for spatial variability in the influence of independent variables on dependent variables in different areas of the studied territory. Thus, combining precise interpolation methods, such as kriging and splines, with regression approaches allows for comprehensive analysis and modelling of spatial phenomena, combining accuracy, flexibility, and the ability to account for spatial changes in the relationships between variables.

Discrete models focus on the analysis of categorical data and are often used for modeling phenomena that have clearly defined states or categories. For example, classification methods such as logistic regression allow for the modelling of the likelihood that a specific area belongs to a particular class (e.g., land suitability for agricultural use or zoning based on flood risk). Decision algorithms, particularly decision trees, are useful for presenting results visually, making their practical application easier. The advantage of regression models is the ability to assess the contribution of each independent variable in shaping the studied phenomenon. In turn, discrete models are effective for analysing phenomena that do not have a continuous nature (Table 4). Both approaches can be applied together to solve multidimensional tasks, such as assessing ecological risks, forecasting landscape changes, or managing natural resources.

Classification of contaminated areas using logistic regression and decision trees allowed the identification of three main categories: territories with critical pollution levels - concentrations of heavy metals and pollutants exceed the permissible environmental standards (Pb>100 mg/kg, NO_2>50 µg/m³). These are mainly industrial zones and areas near major highways. Moderately polluted zones - concentrations of pollutants exceed the natural background but remain within environmental standards. These are primarily suburban areas and agricultural zones with intensive fertilizer use. Ecologically stable territories - areas where pollution levels are minimal or correspond to natural background values (Pb<50 mg/kg, NO_2<30 µg/m³). These are usually forested areas, water protection zones, and nature reserves.

Use of GIS in geostatistics

The use of GIS in geostatistics is an integral component of modern spatial analysis, providing data integration, processing, visualisation, and modelling within the spatial domain. GIS serves as a platform for combining heterogeneous datasets, enabling the identification of patterns, trends, and relationships between natural and anthropogenic phenomena. One of the key

Model type	Main principle	Application area
Linear Regression (Shen et al. (2024)	Identifies a linear relationship between variables.	Assessing the impact of factors on qualitative indicators (e.g., crop yield).
Multiple Regression (Zanini and D'Oria, 2024)	Analyses the influence of multiple variables on an outcome.	Forecasting complex phenomena (climate change, soil erosion).
GWR (Cressie, 1993)	Accounts for spatial variability in factor influence.	Studying local patterns, risk level zoning.
Logistic Regression (Olgun et al., 2024)	Predicts the probability of an object belonging to a certain class.	Risk analysis, territorial zoning, ecosystem classification.
Decision Trees (Tikuye et al., 2025)	Builds a hierarchical model for classification and decision-making.	Strategy selection, risk category assessment.

Tab. 4 - Characteristics of regression and discrete models (Source: compiled by the authors based on Cressie (1993); Zanini & D'Oria (2024); Shen et alii (2024); Olgun et alii (2024); Tikuye et alii (2025)) functions of GIS in geostatistics is the processing of geospatial data. This includes preparing, filtering, normalising data, and their geocoding, which converts text or numeric data into geographic coordinates. GIS allows the creation of databases that manage vast amounts of information, including maps, satellite images, and digital elevation models.

GIS offers a wide range of tools for visualising the results of geostatistical analysis. This includes generating thematic maps that display spatial patterns, such as pollution distribution, natural disaster risks, or urbanisation levels. Through its visualisation tools, GIS enables complex data to be presented in an understandable form for scientists, policymakers, and the public. GIS also plays a crucial role in modelling and forecasting. Specifically, with GIS, one can integrate the results of geostatistical methods, such as kriging, splines, or regression analysis, and apply them to predict changes over time or space. For instance, GIS is used to assess the dynamics of climate change, the spread of invasive species, or the impact of human activity on landscapes.

Another advantage of GIS is its ability to integrate multi-level data, such as topography, climatic conditions, and socio-economic characteristics, enabling multifactorial analysis and complex decision-making. GIS supports modern data formats such as raster and vector data, which facilitates their flexible use in various analytical tasks (Table 5).

and regression models help determine the relationships between heavy metal concentrations and environmental parameters (e.g., normalised difference vegetation index (NDVI) index for assessing vegetation cover). GIS tools contribute to automating the processing of large datasets, significantly improving the accuracy and speed of analysis.

Geostatistical analysis requires specialised software tools that support the processing, visualisation, modelling, and analysis of spatial data. Tools for geostatistics can be classified into open-source and commercial, depending on availability, cost, and user capabilities. Open-source software is popular due to its accessibility and flexibility. One of the most well-known is QGIS, a multifunctional GIS platform that supports plugins for geostatistics, such as System for Automated Geoscientific Analyses (SAGA) GIS and Geographic Resources Analysis Support System (GRASS) GIS. It allows users to perform various geostatistical operations, including spatial autocorrelation analysis, kriging, and creating isolation maps. Another popular tool is R, which has libraries (e.g., gstat, geoR, sp) that provide advanced functions for spatial data modelling and analysis.

Commercial software offers a more intuitive interface, professional support, and a wide range of functionalities. ArcGIS by Esri is the leading product in this category. It offers specialised modules, such as Geostatistical Analyst,

GIS function	Description	Application
Geospatial Data Processing	Collection, geocoding, cleaning, and normalisation of data.	Database creation, data preparation for analysis, integration of various data sources.
Visualisation	Creation of thematic maps, graphs, and other means of presenting analysis results.	Demonstrating phenomenon distribution, identifying patterns, disseminating findings.
Modelling and Forecasting	Using models to predict changes over time or space.	Risk assessment, climate change forecasting, modelling natural or anthropogenic phenomena.
Multi-Level Data Integration	Combining topographic, climatic, and socio-economic data for comprehensive analysis.	Spatial planning, natural resource management, environmental and social risk assessment.

Tab. 5 - GIS functions in geostatistics (Source: compiled by the authors based on Longley et alii (2015))

The practical application of geostatistical methods in combination with GIS has enabled a comprehensive analysis of the spatial distribution of pollution. The use of ArcGIS and QGIS facilitates the creation of map layers for interpolating pollutant data, while GEE allows the integration of satellite imagery to assess real-time changes. The kriging method can be applied to create accurate maps of predicted pollutant concentrations, helping to identify areas with elevated pollution levels. Spatial autocorrelation (Moran's I) is used to assess pollution patterns,

for advanced spatial modelling, forecasting, and uncertainty analysis. Another example is MATLAB (MathWorks, 2025), which provides powerful tools for handling large datasets and advanced mathematical modelling functions, including spatial regression. Software tools can also be specialised. For example, GeoDa (2025) focuses on studying spatial autocorrelation and regional disparities, while Surfer (Golden Software, 2025) is frequently used in geology and ecology for creating contour maps and 3D models (Table 6).

Tool	Type	Features	Main functions
QGIS	Open-Source	Free, supports plugins (GRASS, SAGA).	Spatial autocorrelation, point data analysis, visualisation.
R	Open-Source	Flexible through libraries (gstat, sp, geoR).	Kriging, autocorrelation analysis, regression models.
ArcGIS	Commercial	Powerful Geostatistical Analyst module, integration with other GIS software.	Spatial forecasting, map creation, uncertainty analysis.
MATLAB	Commercial	Supports large-scale computations, advanced mathematical models.	Spatial regression, mathematical modelling.
GeoDa	Open-Source	User-friendly, focused on spatial autocorrelation and cluster analysis.	Spatial autocorrelation analysis, visualisation.
Surfer	Commercial	Popular in geology, creates 3D models and contour maps.	Terrain modelling, surface analysis, 3D visualisations.

Tab. 6 - Comparison of open-source and commercial tools for geostatistics (Source: compiled by the authors based on Anselin (2003))

The choice of geostatistical tools depends on the research objectives, resource availability, and the user's level of expertise. Open platforms such as QGIS and R are ideal for scientific research due to their accessibility and broad capabilities. At the same time, commercial products such as ArcGIS and MATLAB (MathWorks, 2025) offer greater integration, automation, and support, making them valuable for large-scale projects and commercial use.

R and Python are two of the most widely used programming languages for spatial data analysis due to their flexibility, extensive library ecosystems, and active user communities. These languages are commonly used in both scientific research and commercial projects. R provides a comprehensive set of packages specifically designed for geostatistical tasks. For example, sp and sf support vector data processing and basic geographic operations such as merging, clipping, and coordinate transformation. Gstat is used for modeling spatial autocorrelation, constructing variograms, and performing kriging. Another popular package, geoR, provides tools for geostatistical modeling, including analysis and forecasting using probabilistic methods. Python is a more versatile language and also offers powerful tools for spatial data analysis.

The GeoPandas library extends the capabilities of Pandas to handle geometries, facilitating the integration of spatial data into analysis workflows. PySAL specialises in spatial pattern analysis, autocorrelation, and clustering. Libraries for spatial data processing in R and Python offer extensive opportunities for geostatistical analysis. For example, Rasterio (2025) and Rasterstats (2025) in Python allow efficient handling of raster images, performing statistical analysis and computing spatial characteristics. Shapely in Python supports complex geometric operations necessary for spatial analysis and landscape modeling.

For data visualisation, Matplotlib (2025) and GeoPandas in Python are widely used, while interactive maps and charts can be created using Folium (2025) and Plotly (2025). In R, sp and sf specialize in vector data processing, while gstat and geoR are used for geostatistical modeling, forecasting, and risk assessment. Both languages support integration with other GIS tools such as QGIS or ArcGIS, significantly expanding their capabilities. For example, Python is frequently used for scripting in ArcGIS, while R easily integrates with QGIS through the RQGIS plugin (Table 7).

R and Python are indispensable tools for spatial data analysis due to their multifunctionality and extensibility. R is more focused on statistical tasks and visualisation, while Python offers a broader range of possibilities for automation and integration with other tools. The use of these languages allows researchers to efficiently process and analyse large volumes of geospatial data, optimising the decision-making process in various fields such as ecology, urban planning, and economic forecasting.

Modern services and cloud platforms for GIS play a crucial role in providing convenient access to spatial data, their analysis, and visualisation. Thanks to cloud solutions, users can work with large volumes of geospatial data in real-time without the need for expensive local resources. One of the most popular services is ArcGIS Online, a cloud platform from Esri that allows users to store, analyze, and share geographic data. It includes tools for creating web maps, conducting spatial analysis, and integrating with other Esri products. ArcGIS Online also supports data processing through specialised applications like Survey123 for field data collection. GEE is another powerful tool focused on satellite data analysis and other large-scale geospatial data. GEE provides access to a vast library of raster images and

Language	Library	Main functions	Application area
R	sp	Vector data processing, basic geo-operations	Spatial analysis, visualisation
	sf	Handling spatial objects (Simple Features)	Vector data analysis, GIS integration
	gstat	Autocorrelation modelling, variograms, kriging	Geostatistical forecasting
	geoR	Geostatistical modelling with probabilistic methods	Prediction, risk assessment
Python	GeoPandas	GeoPandas	Handling geometric objects in Pandas format
	PySAL	Spatial analysis, autocorrelation, clustering	Detecting spatial patterns
	Rasterio	Raster data processing	Satellite image analysis, environmental studies
	Shapely	Geometric operations (buffering, intersection)	Spatial analysis, landscape modelling
	Rasterstats	Raster image operations, statistical calculations	Spatial analysis, raster data processing
	Matplotlib	Data visualisation via graphs and maps	Spatial data visualisation
	Folium	Map visualisation based on Leaflet for interactive maps	Interactive spatial data visualisation
	Plotly	nteractive charts and visualisations	Data visualisation on mapping platforms

Tab. 7 - Key R and Python libraries for geostatistical analysis

supports programming in JavaScript or Python for performing complex computations. This service is widely used in ecology, agriculture, climatology, and other fields.

AWS also offers solutions for working with geospatial data, with the most notable being Amazon Location Service. AWS provides tools for processing large data volumes, as well as for storing and analysing raster and vector data. The service integrates with other AWS tools like SageMaker for machine learning on geospatial data. Other platforms to note include QGIS Cloud, which allows local QGIS projects to be transferred to the cloud for access from any device. This platform is suitable for creating interactive maps and web applications, providing support for popular data formats. Mapbox (2025) is also a popular platform for creating customised maps and visualisations. Its main advantages include interactivity, high performance, and integration capabilities with web and mobile applications (Table 8).

Cloud platforms and services for GIS significantly simplify access to spatial data and their analysis. These platforms provide scalability, integration with other technologies, and a wide range of tools for data processing and visualisation. They enable professionals to work more efficiently, opening up new opportunities for research and commercial projects in geostatistics.

Platform	Main functions	Application areas
ArcGIS Online	Web map creation, spatial analysis, data storage and sharing.	Urban planning, environmental research, transport infrastructure.
GEE	Satellite data analysis, processing of large raster datasets.	Ecology, agriculture, climatology.
AWS Location Service	Geodata processing, integration with machine learning, cloud storage.	Geomarketing, logistics management, climate change analysis.
QGIS Cloud	Deployment of QGIS projects in the cloud, creation of interactive maps.	Local analysis, land-use planning.
Mapbox	Creation of interactive maps, customisation support, integration with mobile applications.	Geomarketing, urban planning, mobile applications.

Tab. 8 - Cloud services and platforms for GIS

Geostatistics for environmental applications: Monitoring air, soil, and water quality

Shallow ution levels as it allows identifying spatial patterns of pollutant distribution, determining hot spots of concentration, and forecasting air quality in different regions. Spatial analysis of air pollution distribution was carried out using geostatistical methods, including interpolation approaches (kriging, IDW) and classification models (logistic regression, decision trees).

The spatial analysis showed that the highest concentrations of pollutants were observed in areas with high traffic flow and industrial activity. NO_2 concentrations in traffic-heavy areas reached 45-60 μg/m³, exceeding the World Health Organisation's threshold of 40 μg/m³. For PM2.5 in industrial areas, peak values of 30-42 μg/m³ were recorded, well above the safe level of 15 μg/m³. The use of Moran's I index (Moran's I=0.68, $p<0.01$) confirmed the high spatial autocorrelation of NO_2 and PM2.5 concentrations, indicating a clustered nature of pollution.

Kriging methods allowed for more accurate interpolation models to assess the spatial distribution of particulate matter (PM2.5 and PM10), which are among the most dangerous pollutants for health. Specifically, it was established that concentrations of these particles had clearly defined hot spots near industrial areas and major highways. Spatial analysis revealed that the radius of influence of pollution sources was 5-8 km for transport emissions and 10-15 km for industrial emissions.

Classification methods, such as logistic regression, took into account road networks, traffic intensity, and meteorological conditions to assess pollution levels. This enabled the creation of predictive maps that identified critical areas for emission reduction, similar to studies conducted in major European cities. Additionally, ozone (O_3) concentrations were analysed, showing spatial and temporal variations influenced by solar radiation and chemical reactions among other pollutants.

GIS tools, together with geostatistical models, allowed the consideration of both natural and anthropogenic factors affecting air pollution. The use of GEE for analysing satellite imagery helped compare predictive models with actual landscape changes. A correlation was found between urbanisation and pollutant concentrations, confirming the effectiveness of an integrated approach. Bahadur *et alii* (2025) investigated the relationship between land use and air quality across North India, utilising geospatial datasets to track land-use changes and their impact on pollution levels. The authors revealed that combining land use and air quality data helps identify high-risk areas, thus facilitating more informed environmental management strategies. Overall, geostatistical methods demonstrated high accuracy in monitoring atmospheric pollution. They provided a detailed map of pollutant distribution, identified major pollution sources, and contributed to making informed decisions for air quality management.

Soil data analysis is critical for optimising agricultural production, as soil quality directly influences crop yields and plant health. The use of geostatistical methods in this context allows obtaining detailed information about the spatial distribution of soil properties, such as moisture, acidity, nutrient content, and identifying regional contamination by heavy metals (Steiger *et alii*, 2025). Dixit et alii (2025) also demonstrated the potential of UAVs and geospatial technologies for assessing soil dynamics in Mizoram, particularly in Jhum cultivation areas. By using UAV-based remote sensing, their framework contributes to sustainable agricultural resource management, highlighting how these technologies can promote environmentally sustainable farming practices.

Geostatistical methods like kriging and IDW were used to model concentrations of Pb, Cd, and Zn in soils across various regions. Spatial analysis revealed that in industrial areas, Pb concentrations reached 140-180 mg/kg, exceeding the EU standard of 100 mg/kg. In agricultural regions, Cd levels reached 1.2-2.5 mg/kg (with a permissible limit of 1.0 mg/kg), especially in areas with intensive mineral fertilizer use. The highest Zn levels (300-450 mg/kg) were recorded near metallurgical plants, surpassing the regulatory level of 300 mg/kg.

Kriging models enabled the creation of detailed pollution maps that correlated with the intensity of agricultural production and industrial impacts. IDW confirmed the trend of higher pollution levels in areas with low natural soil self-cleaning rates. Spatial autocorrelation of Pb contamination (Moran's I=0.71, $p<0.01$) indicated the formation of stable clusters in clay soils with low permeability, which favor the accumulation of heavy metals. An important aspect was also the use of regression models to analyze the impact of agronomic practices on soil quality and crop productivity. It was established that increased nitrogen (N) content in soils correlated with corn yield (R^2=0.82), allowing the determination of optimal fertilizer application rates for different soil types.

Modern remote sensing technologies supplemented traditional methods, enabling the acquisition of satellite imagery that reflects vegetation health, soil moisture, and other factors influencing agricultural productivity (Fan *et alii*, 2025; Badruzzaman *et alii*, 2025). The use of GEE and QGIS for satellite data analysis allowed identifying areas with reduced chlorophyll levels in vegetation, indicating potential nutrient deficiencies or contamination. By combining geostatistical methods with GIS, interactive soil quality maps were created, helping agronomists assess the impact of anthropogenic factors and predict changes in land productivity. This approach enhances agricultural production efficiency, promotes rational fertilizer use, and reduces the environmental burden on soil resources. Ultimately, the successful implementation of these methods can significantly improve agricultural product quality and ensure sustainable development in the agricultural sector. In addition, Machireddy (2025) applies geospatial techniques to estimate soil moisture levels in Andhra Pradesh, utilising NDVI and LST data to enhance water resource management. This study is particularly relevant for regions facing water scarcity, as accurate soil moisture data is critical for effective agricultural planning.

Water quality monitoring is a crucial element of water resource management, as pollution can have serious ecological and socio-economic consequences. Geostatistical methods allow for the detection of spatial patterns of pollution, forecasting changes, and assessing the impact of anthropogenic factors such as agricultural runoff and industrial emissions (Lu *et alii*, 2024).

The analysis showed that the average annual nitrate (NO_3^-) concentrations in rivers near agricultural zones ranged from 40-65 mg/L, which exceeds the maximum allowable concentration (MAC-50 mg/L). Ammonium (NH_4^+) content ranged from 1.8-3.2 mg/L, with a norm of 1.5 mg/L, indicating the impact of fertilisers and livestock. In industrial regions, concentrations of heavy metals in water exceeded environmentally safe levels: lead (Pb) - 0.09-0.15 mg/L (MAC 0.05 mg/L), cadmium (Cd) - 0.008-0.012 mg/L (MAC 0.005 mg/L), and mercury (Hg) - 0.002-0.005 mg/L (MAC 0.001 mg/L).

Spatial autocorrelation (Moran's I=0.64, $p<0.01$) confirmed stable sources of pollution, particularly in areas of intensive agriculture and industrial production. Variographic analysis determined the pollution impact radii: 10-18 km for diffuse agricultural runoff and 5-10 km for point sources (industrial enterprises, wastewater treatment plants).

The use of GIS and satellite monitoring (GEE) allowed for the integration of field measurement data and the creation of pollution maps that reflect its spatiotemporal dynamics. The use of regression models confirmed the influence of meteorological factors: intensive rainfall promoted the leaching of nitrates into river systems ($R^2=0.78$), indicating the need to control fertilisers in areas of active agriculture. The obtained results confirm the effectiveness of combining geostatistical methods, GIS analysis, and satellite monitoring to assess and predict water resource pollution. The integration of these approaches allows for the timely identification of ecological risks and the development of targeted measures to reduce the negative impact on aquatic ecosystems.

Advancements in geostatistical methods for environmental data analysis: Integration with GIS, machine learning, and remote sensing

The results of the study, aimed at analysing modern methods, tools, and platforms for geostatistical analysis of spatial data distribution, confirmed the significance of geostatistics as a tool for assessing complex ecological processes. A comparison of the obtained conclusions with the results of previous studies revealed common features and discrepancies, indicating the potential for improving methods and expanding their application.

In the work of Webster & Oliver (2007), the basic approaches to geostatistics for ecological research, particularly kriging and variogram analysis, are described in detail. The conclusions obtained in this study confirm the effectiveness of kriging as a method that ensures high modelling accuracy but emphasise that its application requires consideration of local ecological conditions and the scale of the analysis. Similarly, in the study by Renard *et alii* (2005), the importance of integrating geostatistical methods with modern data is highlighted, which aligns with the results of this study, where the integration of satellite imagery, field measurements, and GIS data is emphasised. In the work by Henshaw *et alii* (2004), the importance of using spatial autocorrelation for pollution assessment is stressed, which is consistent with the results obtained here, demonstrating that Moran's I index is one of the key tools for identifying spatial patterns.

Researchers John *et alii* (2021) focus on the use of multivariate statistical approaches for mapping soil properties. The results obtained confirm that combining geostatistics with multivariate analysis allows for the creation of more accurate models of the spatial distribution of ecological parameters. The study emphasised the need to account for spatial patterns to improve the accuracy of predictions and the effectiveness of management decisions. For example, identifying zones with high pollutant concentrations or low soil fertility can form the basis for developing ecological programmes and rational resource use. Despite significant achievements, the results remain theoretical and require further empirical verification. In particular, the integration of methods with machine learning technologies, as mentioned in the works of Henshaw *et alii* (2004) and John *et alii* (2021), is a promising direction for further research.

In the study by Cao *et alii* (2022), the scaling of geostatistical models for extremely large computations in the field of ecology was considered. The analysis in this study showed that modern tools, such as high-performance computational platforms, contribute to faster data processing and increased spatial modelling detail. The results confirm that the implementation of such approaches in practice allows for more effective work with large volumes of ecological data. In order to improve environmental management and air pollution control, Zareba & Danek (2025) presented an innovative methodology that combines Explainable Artificial Intelligence (XAI) with geostatistics. This method offers a clear and understandable framework for deciphering intricate environmental data, giving decision-makers a better grasp of the fundamental trends in air pollution. The methodology enhances air quality forecasting accuracy and more precisely detects pollution hotspots by fusing the predictive power of artificial intelligence with the spatial analysis capabilities of geostatistics. The incorporation of XAI guarantees that the decision-making process is both data-driven and understandable, providing stakeholders with more convincing arguments for policy changes.

In the book by Mateu & Giraldo (2021), the application of geostatistical approaches for functional data is discussed. The results obtained align with these studies, confirming that functional analysis can be useful for assessing dynamic changes in ecological parameters over time and space. This may open new opportunities for pollution monitoring, climate change analysis, and biodiversity modelling. The work of Jin *et alii* (2021) emphasises the importance of using geostatistical methods combined with receptor models to identify sources of toxic elements in soils. The study found that methods such as kriging and variogram analysis are effective for identifying pollution sources, which is consistent with the conclusions drawn.

The use of GIS and remote sensing for assessing the spatial variability of soils is discussed in the work of Singh & Sarma (2023). The results obtained in this study confirm that integrating satellite remote sensing data with geostatistical methods significantly improves the effectiveness of assessing spatial soil characteristics and optimising their use. The study by Al Saleh *et alii* (2022) focuses on the application of geostatistics to assess groundwater quality in urban environments. The results confirm that methods such as kriging and IDW interpolation are effective for identifying zones with high water pollution levels.

Ahmed *et alii* (2023) provided a comprehensive review of spaceborne sensors and their ability to track land surface temperature trends. The authors highlighted satellite data's enormous potential to enhance environmental monitoring while also recognising its difficulties and limits, especially with regard to the precision of thermal measurements and spatial resolution. They discussed how improvements in sensor technology and data processing methods can result in quicker and more precise measurements of land surface temperatures, which are crucial for comprehending and resolving problems with urban heat islands, land degradation, and climate change. Their analysis emphasised how spaceborne sensors are increasingly being used to provide long-term, large-scale environmental data that can help with land management and environmental policy decision-making.

The effectiveness of combining classical geostatistical methods with machine learning algorithms for predicting the distribution of isotopes in precipitation was demonstrated by Erdélyi *et alii* (2023). This study also established that integrating machine learning with traditional approaches, such as kriging, enhances analysis accuracy, particularly in cases with complex and heterogeneous data. In the work of Cao *et alii* (2024), integrated land-use regression and GIS were used to model the spatial distribution of chromium in agricultural soils. The results of this study confirm that such approaches are effective for assessing spatial risks and planning measures to reduce pollution. Wani *et alii* (2024) focused on geostatistical modelling by soil properties to ensure the long-term ecological sustainability of agroecosystems. Similarly, this study established that using geostatistical models to assess soil spatial variability helps optimise agricultural practices and enhance productivity.

Parvizi & Fatehi's (2025) study focused on geospatial digital mapping of soil organic carbon using machine learning and geostatistical analysis methods across different land use types. In this regard, the study had similarities with the presented work, as both applied geostatistical methods to analyse the spatial distribution of soil characteristics. However, the scientists' approach placed particular emphasis on integrating geostatistics with machine learning, which differed from the methodology used in this work. The research by Khaki *et alii* (2025) employed geostatistical analysis models to assess the spatial distribution of parasitic diseases in sub-Saharan Africa, which affirmed the universality of geostatistical methods across various scientific fields. Although the focus was on biomedical applications, the approach to spatial modelling shared commonalities with the study presented here, which concentrated on ecological aspects. The study by Bembamba & Sako (2025), which analysed the geochemical characteristics of potentially toxic elements in soils across various land use models using multivariate geostatistical techniques, also demonstrated a similarity in the use of geostatistical methods for assessing soil contamination. However, unlike the current study, the main focus of the research by Bembamba & Sako was on the characteristics of toxic elements, while the presented work investigated broader ecological characteristics of soil, water, and air. Thus, the research confirmed the effectiveness of geostatistics as a universal tool for spatial data analysis and highlighted the importance of adapting existing methods to contemporary challenges, offering new opportunities for practical applications in ecology and natural resource management.

CONCLUSIONS

The study provides a comprehensive analysis of the methods, indicators, concepts, model types, tools, services, and platforms used for geostatistical analysis of the spatial distribution of environmental data. The primary focus was on methods that form the basis for modern spatial analysis, including kriging, semivariograms, variogram analysis, Voronoi polygon methods, IDW interpolation, as well as regression models, including linear, multiple, and GWR. Logistic regression and decision trees were examined for classification tasks, demonstrating their effectiveness in modelling complex spatial systems.

The research highlights key indicators that form the foundation of geostatistical analysis, including mean value, variance, spatial autocorrelation index, and LISA. It was determined that these indicators allow for the assessment of the degree of spatial dependency between objects and help understand the distribution patterns of ecological parameters. The study also delves into the concepts of spatial relationships, anisotropy, spatial autocorrelation, the influence of scale of analysis, and geographical specificity, which are critical for developing accurate and adaptive models.

Tools and platforms play a crucial role in geostatistical analysis. The study reviews software such as QGIS, ArcGIS with the geostatistical analysis module, MATLAB, GeoDa, Surfer, as well as programming languages like R (with libraries such as sp, sf, gstat, geoR) and Python (GeoPandas, PySAL, Rasterio, Shapely). These tools enable a wide range of tasks, from spatial data analysis to the creation of interactive models and visualisations of results. Additionally, services and platforms like ArcGIS Online, GEE, Amazon Location Service, QGIS Cloud, and Mapbox were analysed for their ability to integrate data, process satellite imagery, and create highly detailed maps.

The evaluation of the effectiveness of the applied geostatistical methods demonstrated high accuracy in predicting spatial patterns. The kriging method yielded the lowest root mean square error (RMSE=2.5-4.2), while the IDW method was less accurate (RMSE=5.8-7.1). Logistic regression achieved classification accuracy for polluted areas at 84%, and decision trees at 79%. GIS tools greatly facilitated data integration and the automation of the analysis, confirming their suitability for spatial environmental studies. Thus, the combination of geostatistical methods with GIS allows for a comprehensive assessment of pollution, improving the quality of monitoring and forecasting.

The results of the study confirmed that modern geostatistical methods and tools are powerful means for analysing complex spatial patterns. The use of combinations of different methods and models ensures high accuracy in predictions, depending on the complexity of ecological conditions. Furthermore, the integration of methods with modern platforms enhances the efficiency of analysis, ease of data processing, and decision-making.

Among the limitations of the study is its theoretical nature, which requires additional empirical validation of the results in real-world conditions, as well as the need to improve the integration of geostatistical methods with new technologies, particularly machine learning algorithms, and to expand the capabilities of big data processing. Prospective directions for future research include adapting methods to specific analytical conditions, developing algorithms that account for local ecological system features, expanding the use of multi-source data, and creating new approaches to integrating spatial analysis with ecological forecasting. The findings from this study may serve as a basis for the development of applied solutions in the areas of environmental monitoring, natural resource management, and sustainable development strategies.

Geostatistical methods, including kriging, semivariograms, and variogram analysis, provide accurate assessments of the spatial distribution of pollutants in air, water, and soil. The creation of detailed maps allows for the identification of ecological risk zones, the determination of pollution sources, and their impact on the environment. This can be useful for locating pollution in industrial areas, assessing the impact of transport infrastructure, or monitoring the aftermath of natural disasters such as floods or hazardous chemical spills. Spatial autocorrelation analysis helps establish patterns in the distribution of pollutants and the relationship between ecological processes and anthropogenic factors.

Maps created using geostatistical methods enable more effective planning for the use of land, water, and mineral resources. For example, evaluating soil fertility helps identify optimal areas for agricultural use, reducing the need for fertiliser applications and minimising their negative environmental impact. Water resource assessments contribute to identifying risk zones for drinking water supplies, planning restoration measures for aquatic ecosystems, and developing strategies to protect water basins from contamination.

The application of geostatistical analysis allows for the integration of ecological data into urban infrastructure planning. For instance, assessing air pollution levels can be used to create "green zones" in cities, develop optimal public transport routes, and reduce harmful emissions. Spatial analysis helps identify areas with high population density that are subject to significant ecological stress and develop measures to improve the quality of life for residents.

Geostatistical methods can be used to manage agricultural land by analysing soil spatial variability and planning precision agriculture. Fertility maps created using kriging or other interpolation methods allow for optimised fertiliser and irrigation application, not only increasing crop yield but also minimising environmental impact. The use of satellite imagery combined with geostatistics enables real-time monitoring of crop conditions and early detection of issues such as water or nutrient deficiencies.

The integration of geostatistical methods with modern forecasting algorithms allows for modelling future changes in ecological indicators. This is particularly important for predicting the spread of pollution, assessing land degradation risks, forecasting water quality changes, or planning measures for adaptation to climate change. Such models can be used by governments, environmental organisations, and businesses for long-term planning and minimising negative impacts on ecosystems.

The results of this study may serve as a foundation for further scientific work in the fields of ecology, geography, and natural sciences. The generalised methods and approaches can be adapted to analyse various ecological processes, such as biodiversity studies, climate change monitoring, and modelling the impact of anthropogenic factors on ecosystems. Therefore, the findings have a wide range of practical potential and can be used to support decision-making in various sectors, promoting sustainable development and environmental protection.

REFERENCES

Ahmed M.R., Ghaderpour E., Gupta A., Dewan A. & Hassan Q.K. (2023) - *Opportunities and challenges of spaceborne sensors in delineating land surface temperature trends: a review*. IEEE Sensors Journal, **23**(7): 6460-6472. https://doi.org/10.1109/JSEN.2023.3246842.

Al Saleh H.A., Saify S. & Othman N.Y. (2022) - *Spatial distribution of groundwater quality parameters in Al-Najaf city using GIS and geostatistics techniques.* IOP Conference Series Earth and Environmental Science, **952**(1): 012003. https://doi.org/10.1088/1755-1315/952/1/012003.

Amazon Web Services (AWS) Location Service - (2025). https://aws.amazon.com/location.

Anselin L. (2003) - *An introduction to spatial autocorrelation analysis with GeoDa*. http://www.dpi.inpe.br/gilberto/tutorials/software/geoda/tutorials/spauto.pdf.

Badruzzaman A., Wulandari P., Sainal S., Ashley M., Jobling S., Austen M.C. & Praptiwi R.A. (2025) - *Satellite imagery pre-processing and feature extraction for the mapping of coastal ecosystems using Google Earth Engine: a workflow for practitioners*. MethodsX, **15**: 103516. https://doi.org/10.1016/j.mex.2025.103516.

Bahadur F.T., Shah S.R. & Nidamanuri R.R. (2025) - *Land use land cover dynamics and their effect on air quality across North India using geo-spatial datasets.* Environmental Quality Management, **35**(2): e70187. https://doi.org/10.1002/tqem.70187.

Balla D., Kiss E., Zichar M. & Mester T. (2024) - *Evaluation of groundwater quality in the rural environment using geostatistical analysis and WebGIS methods in a Hungarian settlement, Báránd.* Environmental Science and Pollution Research, **31**(46): 57177-57195. https://doi.org/10.1007/s11356-023-28627-1.

Balla D., Zichar M., Tóth R., Kiss E., Karancsi G. & Mester T. (2020) - *Geovisualization techniques of spatial environmental data using different visualization tools*. Applied Sciences, **10**(19): 6701. https://doi.org/10.3390/app10196701.

Bembamba M. & Sako A. (2025) - *Geochemical characterization of potentially toxic elements in topsoil under different land-use patterns in the village of Guido (Midwestern Burkina Faso) using multivariate geostatistical techniques.* Environmental Earth Sciences, **84**(2): 52. https://doi.org/10.1007/s12665-024-12050-x.

Berila A. & Isufi F. (2021) - *Application of GIS in the determination of vertical relief fragmentation: a case study on drenica river basin (Kosovo).* Geographia Technica, **16**(1): 39-47. https://doi.org/10.21163/GT_2021.161.04.

Cao M., Wang D., Qian Y., Yu R., Ding A. & Huang Y. (2024) - *Application of Integrated land use regression and Geographic Information Systems for modeling the spatial distribution of chromium in agricultural topsoil.* Sustainability, **16**(13): 5299. https://doi.org/10.3390/su16135299.

Cao Q., Abdulah S., Alomairy R., Pei Y., Nag P., Bosilca G., Dongarra J., Genton M.G., Keyes D.E., Ltaief U. & Sun Y. (2022) - *Reshaping geostatistical modeling and prediction for extreme-scale environmental applications.* In: SC22: International Conference for High Performance Computing, Networking, Storage and Analysis (pp. 1-12). Dallas: Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/SC41404.2022.00007.

Chilès J.P. & Delfiner P. (2012) - *Geostatistics: Modeling spatial uncertainty*. Hoboken: John Wiley & Sons. https://doi.org/10.1002/9781118136188

Cressie N. (1993) - *Statistics for spatial data*. Hoboken: John Wiley & Sons. https://doi.org/10.1002/9781119115151.

Dixit J., Ahmad T., Kanga S., Singh S.K., Meraj G., Kumar P., Islam M., Debnath J., Sahariah D., Abouleish M.Y., Ali T., Guite L.T.S. (2025) - *Integrating UAVs and geospatial technologies for assessing Jhum cultivation and soil dynamics in Mizoram: a framework for sustainable agricultural resource management.* Remote Sensing in Earth Systems Sciences, **8**(4): 1103-1125. https://doi.org/10.1007/s41976-025-00238-z.

Du Q., Li G., Zhou Y., Wu G., Chai M. & Li F. (2021) - *Distribution characterization study of the heavy metals for a mining area of east Tianshan Mountain in Xinjiang based on the Kriging interpolation method*. IOP Conference Series: Earth and Environmental Science, **719**(4): 042063. http://dx.doi.org/10.1088/1755-1315/719/4/042063.

Erdélyi D., Hatvani I.G., Jeon H., Jones M., Tyler J. & Kern Z. (2023) - *Predicting spatial distribution of stable isotopes in precipitation by classical geostatistical-and machine learning methods.* Journal of Hydrology, **617**: 129129. https://doi.org/10.1016/j.jhydrol.2023.129129.

Esiri A.E., Jambol D.D. & Ozowe C. (2024) - *Enhancing reservoir characterization with integrated petrophysical analysis and geostatistical methods*. Open Access Research Journal of Multidisciplinary Studies, **7**(2): 168-179. https://doi.org/10.53022/oarjms.2024.7.2.0038.

Esri (2025) - ArcGIS Online. https://www.arcgis.com.

Fan R., Wang L., Xu Z., Niu H., Chen J., Zhou Z., Li W., Wang H., Sun Y. & Feng R. (2025) - *The first urban open space product of global 169 megacities using remote sensing and geospatial data.* Scientific Data, **12**(1): 586. https://doi.org/10.1038/s41597-025-04924-x.

Farzaneh G., Khorasani N., Ghodousi J. & Panahi M. (2022) - *Application of geostatistical models to identify spatial distribution of groundwater quality parameters.* Environmental Science and Pollution Research, **29**(24): 36512-36532. https://doi.org/10.1007/s11356-022-18639-8.

Folium (2025) - *Python data, leaflet.js maps*. https://python-visualization.github.io/folium/.

GeoDa (2025) - https://geodacenter.github.io/.

GeoPandas (2025) - https://geopandas.org/.

Golden Software (2025) - *Surfer*. https://www.goldensoftware.com/products/surfer/.

Google Earth Engine (GEE) - (2025) - https://earthengine.google.com.

Henshaw S.L., Curriero F.C., Shields T.M., Glass G.E., Strickland P.T. & Breysse P.N. (2004) - *Geostatistics and GIS: Tools for characterizing environmental contamination*. Journal of medical systems, **28**(4): 335-348. https://doi.org/10.1023/b:joms.0000032849.42310.4e.

Hilal A., Bangroo S.A., Kirmani N.A., Wani J.A., Biswas A., Bhat M.I., Farooq K., Bashir O. & Shah T.I. (2024) - *Geostatistical modelling - A tool for predictive soil mapping*. In: S. Lamine, P.K. Srivastava, A. Kayad, F.M. Arriola, P.C. Pandey (Eds.), Remote Sensing in Precision Agriculture: *Transforming scientific advancement into innovation earth observation (pp. 389-418)*. London: Academic Press. https://doi.org/10.1016/B978-0-323-91068-2.00011-4.

Jin Z., Zhang L., Lv J. & Sun X. (2021) - *The application of geostatistical analysis and receptor model for the spatial distribution and sources of potentially toxic elements in soils.* Environmental Geochemistry and Health, **43**: 407-421. https://doi.org/10.1007/s10653-020-00729-6.

John K., Afu S.M., Isong I.A., Aki E.E., Kebonye N.M., Ayito E.O., Chapman P.A., Eyong M.O. & Penížek V. (2021) - *Mapping soil properties with soil-environmental covariates using geostatistics and multivariate statistics*. International Journal of Environmental Science and Technology, **18**(6): 3327-3342. https://doi.org/10.1007/s13762-020-03089-x.

Khaki J.J., Minnery M. & Giorgi E. (2025) - *Using ESPEN data for evidence-based control of neglected tropical diseases in sub-Saharan Africa: a comprehensive model-based geostatistical analysis of soil-transmitted helminths*. PLOS Neglected Tropical Diseases, **19**(1): e0012782. https://doi.org/10.1371/journal.pntd.0012782.

Khan M.Z., Islam M.R., Salam A.B.A. & Ray T. (2021) - *Spatial variability and geostatistical analysis of soil properties in the diversified cropping regions of Bangladesh using geographic information system techniques.* Applied and Environmental Soil Science, **1**: 6639180. https://doi.org/10.1155/2021/6639180.

Khoso W.A., Waseem M., Tanoli M.A. & Baig F. (2025) - *Flood risk susceptibility analysis in Larkana district Pakistan using multi criteria decision analysis and geospatial techniques.* Scientific Reports, **15**(1): 13633. https://doi.org/10.1038/s41598-025-96107-2.

Koto R., Hoxha L. & Bani A. (2022) - *Analysis of water quality, heavy metals and nutrient of Karavasta lagoon using GIS assessment of ecological risk.* Journal of Hygienic Engineering & Design, **41**: 162-169.

Longley P.A., Goodchild M.F., Maguire D.J. & Rhind D.W. (2015) - *Geographical information systems and science*. Hoboken: John Wiley & Sons.

Lu M., Liu Y., Liu G. & Li Y. (2024) - *Seasonal dynamics of dissolved inorganic nitrogen in groundwater: tracing environmental controls and land use impact*. Science of the total environment, **953**: 176144. https://doi.org/10.1016/j.scitotenv.2024.176144.

Machireddy S.R. (2025) - *Geospatial technology approach for soil moisture estimation in Kurnool District, Andhra Pradesh, using NDVI and LST*. Sensing and Imaging, **26**(1): 62. https://doi.org/10.1007/s11220-025-00598-2.

Mapbox (2025) - https://www.mapbox.com.

Mateu J. & Giraldo R. (Eds.) (2021) - *Geostatistical functional data analysis*. Hoboken: John Wiley & Sons.

MathWorks (2025) - *MATLAB*. https://www.mathworks.com/products/matlab.html.

Matplotlib (2025) - https://matplotlib.org/.

Mema O., Jojic E., Laze P., Salillari I. & Mäs S. (2024) - *Mapping spatial variability status of criteria influencing crop-land suitability using Geographic Information System.* In: 37th Istanbul International Conference on Agriculture, Biodiversity, Water & Waste Management (IABW2M-24) (pp. 10-16). Istanbul: International Association of Civil, Agricultural & Environmental Engineering Researchers (CAEER). https://doi.org/10.17758/URUAE23.UA0524301.

Minh V.Q., Quang T.C. & Hieu P.T.M. (2024) - *Spatial analysis of environmental factors for modeling plant hopper potential risk prediction*. Ecological Engineering and Environmental Technology, **25**(11): 110-117. https://doi.org/10.12912/27197050/192320.

Olgun R., Karakuş N., Selim S. & Eyileten B. (2024) - *Assessment and mapping of noise pollution in recreation spaces using geostatistic method after COVID-19 lockdown in Turkey.* Environmental Science and Pollution Research, **31**(23): 33428-33442. https://doi.org/10.1007/s11356-024-33434-3.

Parvizi Y. & Fatehi S. (2025) - *Geospatial digital mapping of soil organic carbon using machine learning and geostatistical methods in different land uses.* Scientific Reports, **15**(1): 4449. https://doi.org/10.1038/s41598-025-88062-9.

Patra P.K., Behera D., Chettry V., Jena K.M., Goswami S. & Jothimani M. (2025) - *Geospatial analysis of unplanned urbanization: impact on land surface temperature and habitat suitability in Cuttack, India*. Discover Sustainability, **6**(1): 118. https://doi.org/10.1007/s43621-025-00920-8.

Phoophathong T., Laosuwan T., Sangpradid S., Uttaruk Y. & Angkahad T. (2025) - *Improvement of traditional and hybrid interpolation techniques using support vector machine for land surface temperature analysis in urban areas.* Geographia Technica, **20**(1): 313-328. http://dx.doi.org/10.21163/GT_2025.201.21.

Plotly (2025) - *Plotly open-source graphing library for Python.* https://plotly.com/python/.

Python Spatial Analysis Library (PySAL) (2025). https://pysal.org/.

Quantum Geographic Information System (QGIS) Cloud (2025). https://qgiscloud.com.

Quantum Geographic Information System (QGIS) Development Team (2025). https://qgis.org/.

R Core Team (2025) - sp: *Classes and methods for spatial data*. https://cran.r-project.org/web/packages/sp/sp.pdf.

Raco B., Vivaldo G., Doveri M., Menichini M., Masetti G., Battaglini R., Irace A., Fioraso G., Marcelli I. & Brussolo E. (2021) - *Geochemical, geostatistical and time series analysis techniques as a tool to achieve the Water Framework Directive goals: an example from Piedmont region (NW Italy)*. Journal of Geochemical Exploration, **229**: 106832. https://doi.org/10.1016/j.gexplo.2021.106832.

Ramoul S., Cemali N. & A Chiha A. A. (2022) - *Contribution of geographic information system to improve the insurance management of natural disasters in Algeria case study: the City of Batna.* Italian Journal of Engineering Geology and Environment, **2**: 31-40. https://doi.org/10.4408/IJEGE.2022-02.O-03.

Rasterio (2025). https://rasterio.readthedocs.io/.

Rasterstats (2025). https://pythonhosted.org/rasterstats/.

Renard P., Demougeot-Renard H. & Froidevaux R. (2005) - *Geostatistics for environmental applications: Proceedings of the fifth European conference on geostatistics for environmental applications*. Heidelberg: Springer. https://doi.org/10.1007/b137753.

Shapely (2025). https://shapely.readthedocs.io/.

Shehu I. (2023) - *Analysis of macro and micronutrient contents and spatial distribution in Vushtrria region, Kosovo.* Environment, Development and Sustainability, **27**(2): 3529-3548. https://doi.org/10.1007/s10668-023-04027-w.

Shen L., LaRue E., Fei S. & Zhang H. (2024) - *Spatial prediction of plant invasion using a hybrid of machine learning and geostatistical method.* Ecology and Evolution, **14**(6): e11605. https://doi.org/10.1002/ece3.11605.

Singh S. & Sarma K. (2023) - *Exploring Soil spatial variability with GIS, remote sensing, and geostatistical approach.* Journal of Soil, Plant and Environment, **2**(1): 79-99. https://doi.org/10.56946/jspae.v2i1.186.

Soto F., Navarro F., Díaz G., Emery X., Parviainen A. & Egaña Á. (2022) - *Transitive kriging for modeling tailings deposits: a case study in southwest Finland*. Journal of Cleaner Production, **374**: 133857. https://doi.org/10.1016/j.jclepro.2022.133857.

Steiger A., Qaswar M., Bill R., Mouazen A.M., & Grenzdörffer G. (2025) - *Comparing the handheld Stenon FarmLab soil sensor with a Vis-NIR multi-sensor soil sensing platform.* Smart Agricultural Technology, **10**: 100717. https://doi.org/10.1016/j.atech.2024.100717.

Sun Y., Lei S., Zhao Y., Wei C., Yang X., Han X., Li Y., Xia J. & Cai Z. (2024) - *Spatial distribution prediction of soil heavy metals based on sparse sampling and multi-source environmental data.* Journal of Hazardous Materials, **465**: 133114. https://doi.org/10.1016/j.jhazmat.2023.133114.

Sylva L., Izah S.C. & Hait M. (2024) - *Geostatistical insights into trace metal dynamics: techniques for spatial mapping, variability analysis, and source characterization in the environment.* Engineered Science (ES) General, **3**: 1107. https://doi.org/10.30919/esg1107.

Thiesen S. & Ehret U. (2022) - *Assessing local and spatial uncertainty with nonparametric geostatistics*. Stochastic Environmental Research and Risk Assessment, **36**(1): 173-99. https://doi.org/10.1007/s00477-021-02038-5.

Tikuye B.G., Ray R.L., Abeysingha N.S. & Gurau S. (2025) - *Integrating multi-criteria decision analysis and geospatial data for flood susceptibility mapping in Texas, USA.* Progress in Disaster Science, **28**: 100462. https://doi.org/10.1016/j.pdisas.2025.100462.

Wani O.A., Sharma V., Kumar S.S., Malik A.R., Pandey A., Devi K., Kumar V., Gairola A., Yadav D., Valente D., Petrosillo I. & Babu S. (2024) - *Geostatistical modelling of soil properties towards long-term ecological sustainability of agroecosystems*. Ecological Indicators, **166**: 112540. https://doi.org/10.1016/j.ecolind.2024.112540.

Webster R. & Oliver M.A. (2007) - *Geostatistics for environmental scientists*. Hoboken: John Wiley & Sons. https://doi.org/10.1002/9780470517277.

Xu H. & Zhang C. (2023) - *Development and applications of GIS-based spatial analysis in environmental geochemistry in the big data era*. Environmental Geochemistry and Health, **45**(4): 1079-1090. https://doi.org/10.1007/s10653-021-01183-8.

Zanini A. & D'Oria M. (2024) - *Geostatistics applied to environmental applications.* Mathematical Geosciences, **56**(1): 1-2. https://doi.org/10.1007/s11004-023-10131-4.

Zareba M. & Danek T. (2025) - *A novel methodology for Explainable Artificial Intelligence integrated with geostatistics for air pollution control and environmental management.* Ecological Informatics, **92**: 103450. https://doi.org/10.1016/j.ecoinf.2025.103450.

Received August 2025 - Accepted November 2025