Comparative Analysis of Geostatistical and Non-Geostatistical Interpolation Methods for Mapping Soil Properties in a Heterogeneous Land-Use System

Orhan Atakan GÜRBÜZ[1*], Gülden GÖK[2]

1 Nigde Vocational School of Social Sciences, Nigde Omer Halisdemir University, Asagi Kayabasi Campus, Suleyman Fethi Avenue, Nigde, Turkey
2 Department of Environmental Engineering, Aksaray University, E-90 Highway 7th km, 68100 Aksaray, Turkey
* E-mail: atakan867@hotmail.com

Abstract

In this study, a mixed-use field that is close to downtown Niğde was chosen to test the performance of interpolation techniques. The traditional descriptive statistics were used to investigate if there were significant differences between the deforested and agricultural parts of the study area. A 3-ha area was chosen with 29 sample locations for analyzing spatial dependence and spatial interpolation. The parameters of soil pH, soil electrical conductivity (EC), organic matter (OM), water content (WC), and oil and grease (OG) were analyzed for 29 sample locations. The spatial autocorrelation was tested based on Moran's I statistics, and positive spatial dependence was the highest for OG. Negative autocorrelation was estimated from soil EC. Also, a non-geostatistical interpolation technique called Inverse Distance Weighting (IDW) was used for the interpolation of soil properties throughout the study area. The performance of IDW was lowest for EC and highest for soil pH, soil oil, and grease. For geostatistical analysis, ordinary kriging (OK) interpolation was used for soil properties. Soil OM showed more precise results in OK than other soil properties. The one way ANOVA showed that the analyzed soil properties did not differ significantly between the agricultural and deforested sections of the study area.

Keywords Digital soil mapping; Spatial autocorrelation; Land-use effects; Spatial prediction; Interpolation accuracy; Soil-physicochemical properties

1. Introduction

Auxiliary tools and significant predictions are needed for land usage and planning sustainable growth. Since artificial surfaces and agricultural areas negatively affect soil quality due to uncontrolled agriculture and inaccurate management policies (Fathizad *et al.* 2020, De Siervi *et al.* 2022, Sharma *et al.* 2024), predictions such as soil quality are important tools for sustainable land decision-making (Abdennour *et al.* 2020, Hu *et al.* 2022, Sun *et al.* 2024). Environmental policymakers and environmental scientists need spatial data for the interpretation of the status quo of environmental conditions (Li & Heap 2014, Yang *et al.* 2024).

Spatial autocorrelation is one of the spatial statistical phenomena that interprets the spatial dependence of any dependent variable (Getis 2010). Spatial autocorrelation analysis is an initial step in spatial statistics to test spatial heterogeneity. Environmental studies use local forms of spatial autocorrelation to identify hot and cold spots related to environmental variables (Zhang *et al.* 2008). Trace elements had been analyzed for local and global spatial autocorrelation statistics in the groundwater systems of the Lower Katari Basin, and regional spatial dependencies of trace elements had been found (Quino Lima *et al.* 2020). Another hotspot study was done by Tepanosyan et al. (2019) in urban soils in Yerevan for Pb, Mo, and Ti, and high-level clusters for Pb concentrations were detected (Tepanosyan *et al.* 2019).

Another perspective on spatial data analysis is spatial interpolation via geostatistical methods. Geostatistical methods, which originated from mining practices, used different branches of natural science and engineering (Oliver & Webster 2014). In soil science, the distributions of environmental indicators such as soil organic carbon (Bhunia *et al.* 2018b), electrical conductivity of soil (Abdennour *et al.* 2020), and soil heavy metals (Liu *et al.* 2011, Lu *et al.* 2012, Ghanbarpour *et al.* 2013).

Previous studies have shown that land-use type and local management preferences can affect soil parameters. For instance, Bakhshandeh et al. (2019) investigated soil samples that were collected from natural forest, arable lands, citrus gardens, and paddy fields in northern Iran. The study was included physicochemical and biological analyses were carried out on 62 soil samples taken from the Sari region of Mazandaran province of Iran. The main purpose of this study was to examine the effects of differences in land use on soil properties, and the interpolation was carried out by kriging methods. The results of the study were concluded that the biological growth index was higher in the natural forest areas among other parts of the study area (Bakhshandeh *et al.* 2019). Another similar study was done by Panday et al. (2019) and was conducted in the Dang region of Nepal from agricultural, agroforestry and grassland areas with 120 soil samples. As a result of the study, it was determined that the soil P was quite different according to the land use (Panday *et al.* 2019).

Recent studies have also emphasized that soil physicochemical properties may be strongly influenced by local anthropogenic activities, site-specific soil conditions. In another study, successive extraction was used to measure and identify the concentrations of Cd, Mn, Ni, and Pb in farming soils near a ceramics company in Nigeria. Furthermore, soil pH and particle size analyses were determined (Emurotu *et al.* 2024). Furthermore, sampling plans can impact the results of estimating unsampled locations. In a study, dissolved organic carbon, dissolved total carbon, dissolved inorganic carbon, and dissolved total nitrogen values of 30 soil samples were collected according to two different sample plans (random and grid). The study area was an 18-ha pasture area in Niğde, Turkey. As a result of the study, it was determined from cross-validation results that the spatial distribution modeling of the samples collected according to the grid plan was less inaccurate than the randomly collected sample points (Gök & Gürbüz 2020).

Based on this literature background, the spatial interpolation techniques and spatial autocorrelation context have been used in the environmental science field. The present study evaluates the performances of IDW and OK for mapping selected soil physicochemical properties in a small heterogeneous land-use system. This study also applies the traditional descriptive statistics and one-way ANOVA test to investigate if there were significant differences between the deforested and agricultural parts of the study area. To assess spatial dependence, a spatial statistic test for global spatial autocorrelation (Moran's I) was used. Finally, IDW and OK methods were used to estimate the values of unsampled locations.

2. Materials and Methods

2.1. Study Area

The study site is in Niğde, which is a small city in the central region of Turkey (Fig. 1). The latitude and longitude of the study area were 37°57'10.24"N and 34°41'18.84"E. A cement factory was located 500 m southwest of the area, and there is continued urban fabric 350 m north. Unused seminatural areas and agricultural areas were located east and west of the study field. Sampling points were chosen from a location where non-irrigated arable land and deforested land were neighbors. During the interview with the landowner, it was learned that part A of the field has been used for growing wheat for 12 years, and only organic fertilizers have been used during cultivation.

2.2. Soil sampling and analysis

A random sample plan was used for the prediction of the field, as can be seen in Fig. 1. During the sampling process, Google Earth Pro and ArcGIS v.10.5 were used to visualize the study area and sampling points. Sampling points were generated randomly via Visual Sample Plan (VSP) v. 7 (Saito & McKenna 2007), and the distances between pairs of points varied from 9 meters to 202 meters. 8 samples were collected from Study Area A and 21 samples from Study Area B (Fig.1). The total area of the field was 3 ha. All soil samples were collected from topsoil (0-20 cm) on March 15, 2018. Before laboratory analysis, all sampled soils were sieved through a 0.05 mm mesh. Soil pH and soil EC were analyzed with a Hach HQ411D digital pH meter. The gravimetric methods used to analyze soil WC and soil organic matter (OM) (Gök & Gürbüz 2020). OG ratios were estimated from Soxhlet extraction (USEPA 1996).

The number of sampling points was determined based on the small size of the study field, sampling accessibility, laboratory constraints and the random sampling design obtained from VSP. Therefore, the 29 sampling locations were considered as the suitable sampling density for the heterogeneous area instead of any optimized sampling threshold for spatial interpolation.

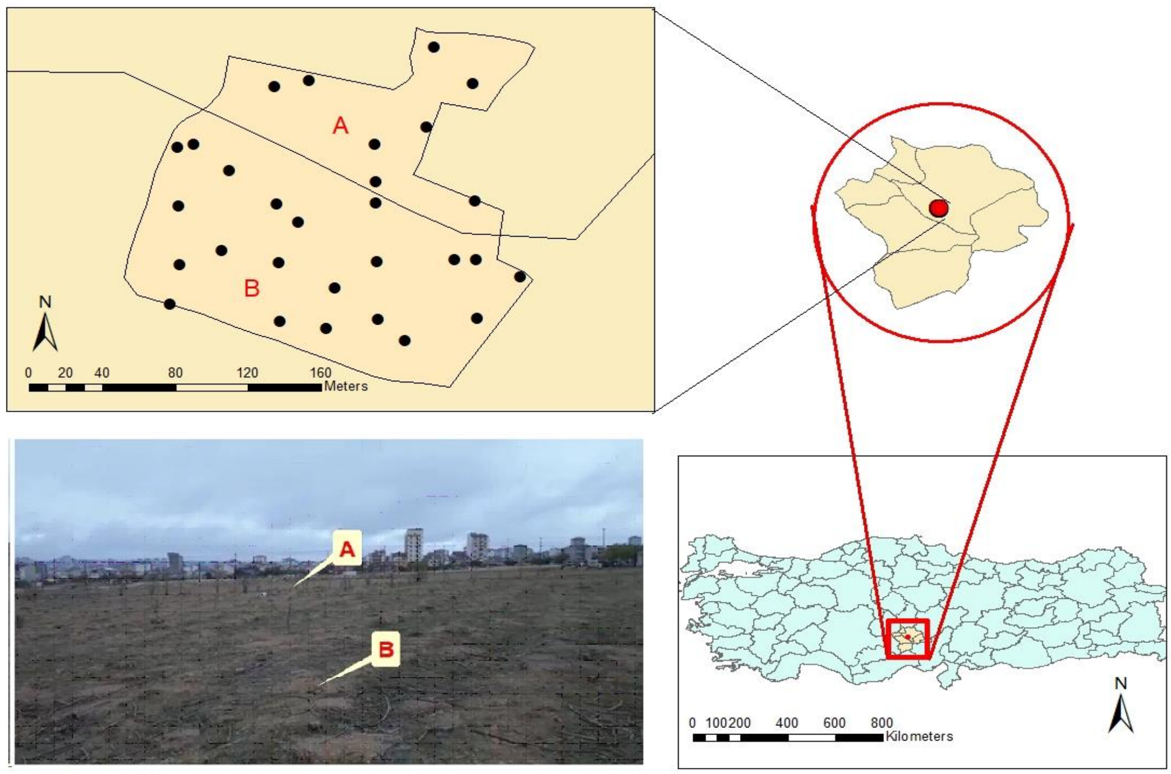

Fig. 1 Study Area, A: Non irrigated arable land, B: Deforested Land

2.3. Statistical Analysis and Spatial Autocorrelation

The descriptive statistical analysis for soil variables was done by SPSS v. 18, and statistical parameters of mean, median, variance, standard deviation, and coefficient of variation were estimated for pH, EC, OM, water content, and OG. Skewness and kurtosis of the parameters were also investigated to control a normal distribution. The coefficient of variation (CV) values were exhibited separation of soil variables (Webster & Oliver 2007). There is a higher dispersion among variables if CV value is greater than 90% (Wei *et al.* 2008). It can be said that a lower variation among variables when CV value is near 10% (Wei *et al.* 2008). Furthermore, the Pearson correlation coefficient was estimated for correlations of soil variables. Also, ANOVA tests were run to evaluate soil variables that any land use change had a statistically important effect on these variables. For statistical analysis, IBM SPSS v.26 was used throughout this study. One of the most widely used techniques for estimating spatial autocorrelation is Moran's I (Eq.1) which is a global index of autocorrelation based on the spatial weights of points or area objects of interest (O'Sullivan & Unwin 2010). The mathematical definition of Moran's I is as follows (Eq.1).

$$I = \left[\frac{n}{(y_i-\bar{y})^2}\right] \left[\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i-\bar{y})(y_j-\bar{y})}{\sum_{i=1}^{n} w_{ij}}\right] \qquad (1)$$

In Eq. 1. the letters i and j represent spatial locations. The variations are estimated based on mean differences multiplied by spatial weights to identify how covariances vary with their spatial locations. Moran's I values range from -1 to +1, where positive values indicate clustering of similar values, negative values indicate spatial dispersion, and values close to zero suggest a random spatial pattern (O'Sullivan & Unwin 2010). The estimated Moran's I is compared with the expected value to determine the degree of spatial autocorrelation (Grekousis 2020). For spatial autocorrelation analysis, ArcGIS v.10.5 was used.

2.4. Interpolation Methods and Data Validation

Inverse Distance Weighting (IDW) is a spatial interpolation method that estimates the variable of an unknown location using inverse distance-based functions (O'Sullivan & Unwin 2010, ESRI 2024). As a default option, the power of the inverse distance function is 2 (ESRI 2024). Rather than automated interpolation, there are statistical ways to estimate variables at unknown locations. Variogram model fitting is essential for precise kriging estimations. The equation of the experimental variogram is as follows:

$$\hat{\gamma}(h) = \frac{1}{2mh}\sum_{i=1}^{m(h)}\{z(x_i) - z(xi + h)\}^2 \qquad (2)$$

In Eq. 2. ,γ.,h. is the spatial lag distance between two locations, h represents the spatial lag, z(xi) and z(xi+h) are sampling locations, and m(h) is the number of sampling location couples at lag h (Oliver

& Webster 2015, Bhunia *et al.* 2018a). The nugget (Co) and sill (Co+C), which are the elements of the semivariogram curve, can be used as an interpretation of spatial dependence (Bhunia *et al.* 2018a). The distribution of soil variables in space was also estimated by ordinary kriging (OK) as a geostatistical analysis. In kriging interpolations, the spatial weights are calculated appropriately, including all sampling locations (O'Sullivan & Unwin 2010). Eq.3. shows the mathematical expression of OK interpolation.

$$Z'(x_0) = \sum_{i=1}^{n} \lambda_i \, Z(x_i) \quad (3)$$

In Eq.3. λ_i represents the kriging weight, $Z'(x_0)$ is the interpolated value of the soil attribute of the unsampled location, and $Z(x_i)$ is the known value of the soil attribute from the sampled location (Chabala *et al.* 2017). To test the accuracy of the interpolation techniques, leave-one-out cross-validation (LOOCV) (Miao *et al.* 2026) was applied along with some validation parameters. Mean error (ME) (Eq. 4.) and root mean square error (RMSE) (Eq. 5.), which are two prominent parameters, were used for evaluating interpolation performances (Szatmári *et al.* 2021). Eq. 4. and Eq.5. where Pi represents predicted values and Oi represents observed values, shows the formulations of ME and RMSE, respectively (Szatmári *et al.* 2021). The Geostatistical Analyst extension of ArcGIS v.10.5 was used for the interpolation of soil variables.

$$ME = \frac{1}{n}\sum_{i=1}^{n}(P_i - O_i) \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2} \quad (5)$$

3. Results

3.1. Statistical Results

Table 1 shows the results of the descriptive statistics of analyzed soil properties. According to these results According to Table 1. the variation of OG is the highest and other variables show moderate variation.

Table 1 Descriptive Statistics of Soil Properties

		N	Mean	Std. Deviation	Minimum	Maximum	Skewness	Kurtosis	CV [%]
pH	Deforested Land	21	8.66	0.10	8.37	8.79			
	Arable Land	8	8.62	0.14	8.37	8.83			
	Total	29	8.65	0.11	8.37	8.83	-0.892	0.623	1.319
EC [µS/cm]	Deforested Land	21	195.66	91.61	127.20	519.00			
	Arable Land	8	291.91	154.97	146.65	626.00			
	Total	29	222.21	117.97	127.20	626.00	2.156	4.795	53.086
WC[%]	Deforested Land	21	2.43	1.32	0.54	5.05			
	Arable Land	8	2.53	1.36	1.04	4.90			
	Total	29	2.45	1.31	0.54	5.05	0.325	-0.961	53.220
OG [%]	Deforested Land	21	0.36	0.56	0.02	2.44			
	Arable Land	8	0.52	0.77	0.01	2.23			
	Total	29	0.40	0.61	0.01	2.44	2.486	5.842	152.413
OM [%]	Deforested Land	21	5.35	1.63	3.04	8.03			
	Arable Land	8	6.55	2.99	4.43	13.18			
	Total	29	5.68	2.10	3.04	13.18	1.697	4.387	37.008

The data on the soil variables were transformed except WC and pH to achieve a normal distribution. Table 2 shows normality test results after transformations of variables. Kolmogorov-Smirnov and Shapiro-Wilk tests were used for analyzing the normal distribution. After transformations, variables reached a statistically significant (Sig. > 0.05) normal distribution.

Table 2 Variable Transformations and Tests of Normality.

		Kolmogorov-Smirnov			Shapiro-Wilk		
	Transformation	Statistic	df	Sig.	Statistic	df	Sig.
pH	None	0.135	29	0.192	0.930	29	0.056
WC[%]	None	0.128	29	0.2	0.949	29	0.174
EC [µS/cm]	Box-Cox (λ=-2.83)	0.121	29	0.2	0.948	29	0.166
OM [g]	Logarithm Base 10	0.135	29	0.190	0.942	29	0.112
OG [g]	Logarithm Base 10	0.099	29	0.2	0.970	29	0.554

Table 3 shows the Pearson correlation coefficient results. According to the results, the soil variables did not have a statistically significant correlation. There were negative and statistically insignificant relationships between OG and other parameters apart from water content. Positive relationships could be seen among other parameters.

Table 3 Pearson Correlation Test Results

		pH	Water Content	OG	OM	EC
pH	Correlation coefficient	1				
	Level of Significance					

Water Content	Correlation coefficient	0.339	1			
	Level of Significance	0.072				
OG	Correlation coefficient	-0.161	0.218	1		
	Level of Significance	0.405	0.257			
OM	Correlation coefficient	0.277	0.030	-0.212	1	
	Level of Significance	0.146	0.878	0.270		
EC	Correlation coefficient	0.309	0.215	-0.244	0.342	1
	Level of Significance	0.103	0.263	0.202	0.069	

According to one-way ANOVA results (Table 4) there wasn't any statistically significant difference in soil variables between study area A and study area B (Fig.1). The mean values of soil variables weren't different based on land use practices.

Table 4 ANOVA results

		Sum of Squares	df	Mean Square	F	Sig.
pH	Between Groups	0.010	1	0.010	0.766	0.389
	Within Groups	0.354	27	0.013		
	Total	0.364	28			
WC	Between Groups	0.058	1	0.058	0.033	0.858
	Within Groups	47.873	27	1.773		
	Total	47.931	28			
OG	Between Groups	0.007	1	0.007	0.020	0.888
	Within Groups	9.410	27	0.349		
	Total	9.417	28			
OM	Between Groups	0.033	1	0.033	1.648	0.210
	Within Groups	0.538	27	0.020		
	Total	0.571	28			
EC	Between Groups	0	1	0	0.528	0.474
	Within Groups	0	27	0		
	Total	0	28			

3.2. Spatial Autocorrelation

Global Moran's I result of variables could be seen in Table 5. Moran's I values are greater than ±0.3 and have strong spatial autocorrelation. According to the results, there was a strong negative spatial autocorrelation for EC. The highest positive spatial autocorrelation values were detected for pH, oil, and grease.

Table 5 Global Moran's I result

	Moran's I	Degree of Autocorrelation
pH	0.135	Moderate
EC	-0.444	Strong
WC	-0.116	Moderate
OM	-0.118	Moderate
OG	0.165	Moderate

3.3. Spatial Interpolations and Data Validations

The first interpolation technique was IDW which is an automated interpolation method without using any statistical methods. Fig. 2 shows the results of IDW interpolation. According to the estimations, higher pH values were located at the center of the study area. WC values showed a dispersed pattern over study area. For EC, a low cluster could be found in the deforested area. There was a cluster of higher values that could be inspected for OG for the northern part of the field.

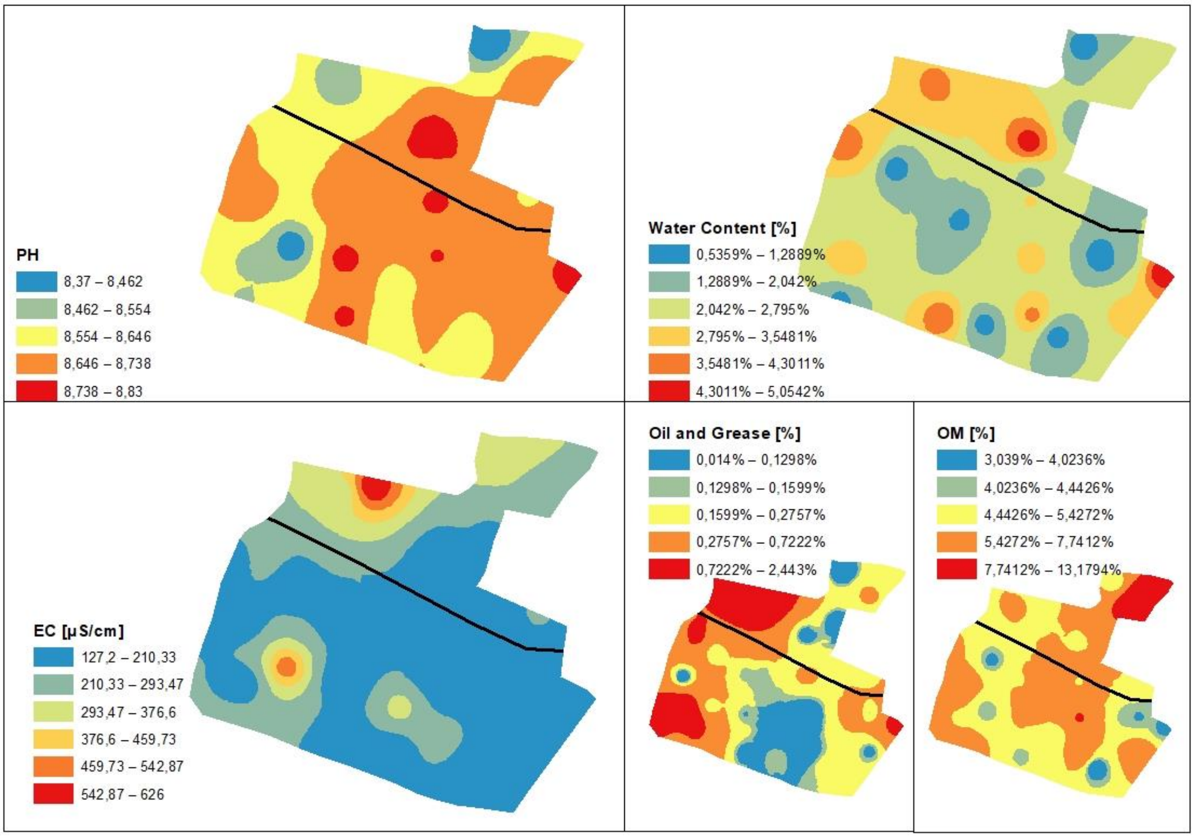

Fig. 1 IDW interpolation of soil properties

Table 6 shows validation parameters for IDW interpolation. As can be seen from table, pH and OG had the lowest error measures. On the other hand, highest error measures were obtained from EC and in this set of materials it can be said that IDW interpolation was not a precise option for interpolation of soil EC.

Table 6 IDW validation measures

Parameters	IDW		
	R^2	ME	RMSE
pH	0.0020	0.0064	0.1211
EC [μS/cm]	0.0843	-3.45236	118.98056
WC [%]	0.0340	-0.04357	1.4835
OM[%]	0.0545	-0.0776	2.4655753
OG [%]	0.0407	-0.00151	0.6254295

The OK estimation results were exhibited more clustered pattern rather than IDW method (Fig.3.). pH values were clustered on east side of the field as a higher cluster. Spatial distribution of WC values was random with moderate scores. There was a high-value cluster on wheat production side of the field for EC and OM values.

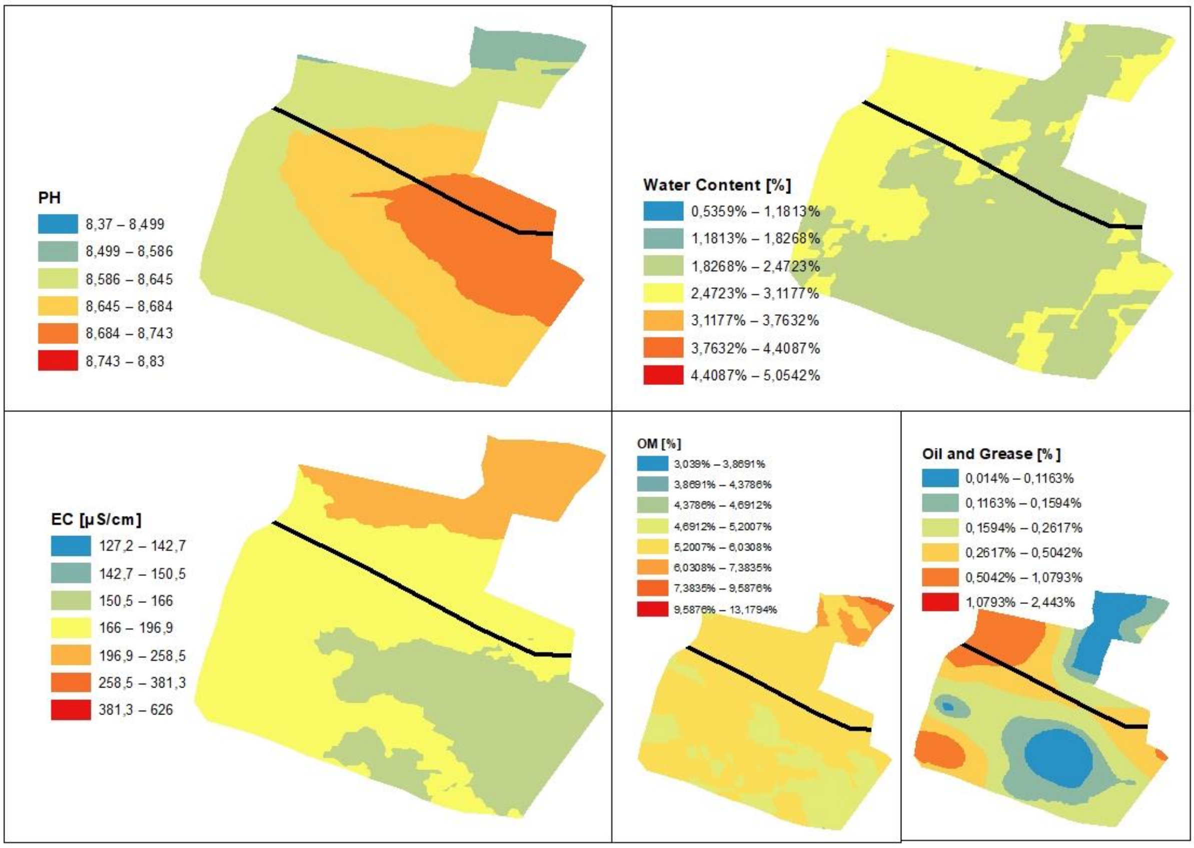

Fig. 2 OK results

In Table 7, semivariogram models and components of semivariogram for each soil property was given. Nugget and sill ratio is another way to assess spatial autocorrelation. The nugget and sill ratio lower than 0.25 can be considered as high degree spatial autocorrelation. The ratio higher than 0.75 low level spatial autocorrelation and the values between 0.25-0.75 are at the moderate level (Bhunia *et al.* 2018a). So, according to the model estimations of semivariogram, the high-level spatial autocorrelation could be seen for OG values. Soil pH of the area of interest had moderate spatial dependence and all other soil parameters low spatial autocorrelation. The difference of the spatial autocorrelation interpolation from Table 4. was the usage of spatial weights. For calculation Moran's I, inverse distance function was used for each soil property as spatial weights. For OK estimation, spatial weights were calculated from semivariogram functions (Oliver & Webster 2015). Although soil pH had lowest error values, the R^2 value of pH for OK was the closest to zero. WCand OM with low spatial autocorrelation degree but had more precise OK estimation than other soil properties The soil EC had the lowest performance for OK.

Table 7 Expected Semivariogram parameters and measures of validation for OK

Parameters	Best Fitted Model	Nugget (Co)	Sill (Co+C)	Partial Sill (C)	Range (m)	Nugget/Sill	R^2	ME	RMSE
pH	Gaussian	0.00959084	0.0158138	0.006223	256.97	0.61	0.0004	0.0015	0.1181
EC	Exponential	1.02E-14	1.24E-14	2.19E-15	112.01984	0.82	0.094	1.759106	110.0436
WC	Circular	1.61495	1.76625	0.1513	137.7104	0.91	0.2898	-0.03379	1.39031
OM	Gaussian	0.0952	0.0983	0.0031	83.896	0.97	0.2411	-0.05848	2.212016
OG	Exponential	0.54179	2.3023	1.7605	256.966	0.24	0.0198	-0.03927	0.608763

4. 4. Discussion

4.1. Variation in soil properties and land use effects

Within this specific localized study, ANOVA test results (Table 4.) revealed no statistically significant differences in the mean values of any soil variables between the agricultural and deforested areas. The absence of statistically significant differences in soil quality parameters may be attributed to the cultivation of wheat in the agricultural area, which is related to stable soil management practices.

The distinct land-use intensities can drastically alter the physical, chemical, and biological balance of the soil. In a Mediterranean landscape in Chile, native forest soils undisturbed by human intervention exhibited the highest Soil Quality Index (SQI = 0.82). When these natural landscapes were converted to agricultural uses, soil quality significantly deteriorated, dropping to an SQI of 0.27 for annual crops, 0.34 for perennial crops, and 0.36 for grassland (Hernández *et al.* 2016). Especially in the Niğde region, the soil quality dramatically changes due to agricultural cultivation for fields where the surface cover is limited. Cultivated lands with limited protective cover (such as potato and bean fields) and degraded pasturelands generated the highest surface runoff and highest erosion rate (Yaşar Korkanç 2018).

In contrast, the agricultural field of the presented study was wheat-planted field; wheat was shown to minimize soil loss and improve soil structure and could be used in crop rotations (Watkins & Lu 1998). Furthermore, wheat provides significant residue cover, which is important for reducing soil erosion and protecting soil quality (Larney *et al.* 2017).

In addition, some studies have reported that differences in land-use types or crop rotation systems do not always lead to statistically significant changes in soil quality parameters. Hoss *et al.* (2018) investigated the long-term impact of different agricultural land-use rotations on soil quality across six locations in Illinois over a 12-year period. The study compared continuous corn, corn-soybean, and corn-corn-soybean rotations by measuring 21 different soil properties. The results demonstrated that the different crop rotations did not affect any of the evaluated soil properties. Shoumik *et al.* (2025) evaluated soil quality across the European Union and the UK under three highly diverse land-use systems: croplands, grasslands, and woodlands, while specific individual soil parameters (such as bulk density, pH, and organic carbon) did show notable differences depending on the land use.

Soil EC, OM, and WC are difficult parameters to interpolate in small heterogeneous fields because they could vary significantly over short distances. EC is strongly affected by soil water content, organic matter, soil texture, and soluble or exchangeable ions, which could complicate spatial prediction under limited sampling density (Kim & Park 2024). Recent geostatistical research also showed that EC can exhibit much higher variability than other soil properties such as pH, and that different soil variables may require different variogram models for spatial prediction (Nogiya *et al.* 2024). Similarly, WC is a highly sensitive soil parameter because it is influenced by micro-topography, slope position, vegetation cover, and soil depth (Zhang *et al.* 2024). Therefore, the relatively weak interpolation performance observed for some variables could not be interpreted only as a limitation of the interpolation method, but also as a reflection of the diverse spatial heterogeneity.

4.2. Comparison of interpolation methods and implications for spatial sampling

In this study, spatial interpolation of soil properties (pH, EC, OM, WC, and OG) was performed using IDW and OK, with the study concluding that OK consistently outperformed IDW by producing lower error measures, while IDW was deemed highly imprecise for parameters like EC. Zhang *et al.* (2015), evaluated 162 spatial datasets and concluded that IDW consistently produces the maximum bias among interpolation methods.

When sample sizes are too small to model a variogram, deterministic methods like IDW are highly recommended over Kriging, as IDW does not rely on measuring spatial structure (Kravchenko 2003). On the other hand, OK was shown to perform well in marine environmental monitoring, reflecting spatial distribution and trends accurately even with small sample sizes (Zou *et al.* 2010).

Furthermore, when applying predictive models to small sample sizes (e.g., 10 to 30 points), simpler methods like Multiple Linear Regression (MLR) are proven to outperform complex machine learning algorithms like Random Forest, which require much larger datasets to stabilize (Schmidinger *et al.* 2024). Also, small sample sizes can be sufficient for spatial interpolation under certain conditions such as adaptive sampling for optimizing the sampling points (Mackman & Allen 2010) and systematic sampling for better interpolation accuracy compared to random sampling (Batur 2022). Sampling density and spatial distribution of sampling points are important factors affecting interpolation accuracy, particularly for highly variable soil properties. In the present study, 29 sampling points were used within a 3-ha field, which may be considered an acceptable sampling density for a site-specific and exploratory comparison of IDW and OK under small-field conditions. However, this number should not be interpreted as an optimum or saturated sampling threshold. İmamoğlu and Sertel (2016) showed that interpolation accuracy in soil moisture mapping could differ depending on the number and spatial distribution of control points. In their study the soil moisture were spatially interpolated from 36 sample points (İmamoğlu & Sertel 2016). Similarly, Safaee et al. (2024) reported that the prediction accuracy of soil properties is influenced by sampling density, sampling distribution, model type, soil property and site-characteristics. In their study it was highlighted that a minimum of 12 to 15 samples for every square kilometer is a necessity for obtaining best predictive models (Safaee *et al.* 2024). Therefore, interpolation performance should not be evaluated only interpolation method, but together with sampling density, sampling design, soil property characteristics, and local land-use heterogeneity. Furthermore, for improved accuracy and reliability, co-kriging could be another option for spatial interpolation of parameters. While OK estimates values at unknown spatial points by only analyzing the spatial auto-correlation of a single primary variable, Co-kriging also incorporates one or more auxiliary variables that are spatially correlated with the primary target with the outcomes of superior statistical performances (Liu & Shang 2026).

5. Conclusion

The ANOVA results indicated that there wasn't any statistically significant difference between the values of the agricultural area and the deforested area of the study field. The spatial autocorrelation results pointed out that the autocorrelation was present negatively for soil EC, soil water content, and soil OM. Positive but less significant autocorrelation was found for soil pH, soil OG. IDW and OK interpolation techniques had low R^2 values. In this set of materials, we could conclude that for more precise results of spatial autocorrelation and interpolation techniques, more sampling locations could be increased. Second, the sampling scheme could be changed to grid sampling with fixed distances between sampling locations for distinct tests of spatial autocorrelation and spatial interpolation.

While the presented study allowed the comparison of IDW and OK under small, restricted study area conditions, the limited sample size could restrict the accuracy of variaogram modeling and the generalization of the modeling results. Future studies should assess distinct sampling sizes and compare random and grid-based sampling schemes to determine a more robust sampling threshold for interpolation techniques. Future studies focusing on a single highly variable soil indicator, such as EC, OM, or WC, with more intensive sampling schemes and auxiliary environmental variables may provide stronger evidence for reducing uncertainty in spatial prediction models in small heterogeneous fields. Also, with the aid of other statistical analyses such as principal component analysis and hierarchical cluster analysis, which could indicate factors that impact the study area with an adequate number of samples.

Acknowledgments

The authors would like to thank Aksaray University and Niğde Ömer Halisdemir University for providing institutional support during the fieldwork and laboratory analysis. The authors also acknowledge the landowner for granting access to the study area.

Funding

This research received no external funding. The study was supported by institutional resources of Aksaray University and Niğde Ömer Halisdemir University.

References

Abdennour, M.A., Douaoui, A., Piccini, C., Pulido, M., Bennacer, A., Bradaï, A., Barrena, J., & Yahiaoui, I., 2020. Predictive mapping of soil electrical conductivity as a Proxy of soil salinity in south-east of Algeria. *Environmental and Sustainability Indicators*, 8, 100087.

Bakhshandeh, E., Hossieni, M., Zeraatpisheh, M., & Francaviglia, R., 2019. Land use change effects on soil quality and biological fertility: A case study in northern Iran. *European Journal of Soil Biology*, 95, 103119.

Batur, M.O., 2022. Assessing spatial interpolation based on sampling size and point geometry in elevation mapping applications. *Journal of Geology, Geography and Geoecology*, 31 (1), 3–9.

Bhunia, G.S., Shit, P.K., & Chattopadhyay, R., 2018a. Assessment of spatial variability of soil properties using geostatistical approach of lateritic soil (West Bengal, India). *Annals of Agrarian Science*, 16 (4), 436–443.

Bhunia, G.S., Shit, P.K., & Maiti, R., 2018b. Comparison of GIS-based interpolation methods for spatial distribution of soil organic carbon (SOC). *Journal of the Saudi Society of Agricultural Sciences*, 17 (2), 114–126.

Chabala, L.M., Mulolwa, A., & Lungu, O., 2017. Application of Ordinary Kriging in Mapping Soil Organic Carbon in Zambia. *Pedosphere*, 27 (2), 338–343.

De Siervi, M., Arreghini, S., & De Iorio, A.F., 2022. Losses of sediment, organic matter, and nutrients in the Argentinean Pampas through rainfall simulation experiments. *Journal of Soils and Sediments*, 22 (9), 2485–2498.

Emurotu, J.E., Azike, E.C., Emurotu, O.M., & Umar, Y.A., 2024. Chemical fractionation and mobility of Cd, Mn, Ni, and Pb in farmland soils near a ceramics company. *Environmental Geochemistry and Health*, 46 (7), 241.

ESRI, 2024. How IDW works—Help | ArcGIS for Desktop [online]. Available from: https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/how-idw-works.htm [Accessed 25 May 2024].

Fathizad, H., Ardakani, M.A.H., Heung, B., Sodaiezadeh, H., Rahmani, A., Fathabadi, A., Scholten, T., & Taghizadeh-Mehrjardi, R., 2020. Spatio-temporal dynamic of soil quality in the central Iranian desert modeled with machine learning and digital soil assessment techniques. *Ecological Indicators*, 118, 106736.

Getis, A., 2010. Spatial Autocorrelation. *In*: M.M. Fischer & A. Getis, eds. *Handbook of Applied Spatial Analysis*. Berlin, Heidelberg: Springer Berlin Heidelberg, 255–278.

Ghanbarpour, M.R., Goorzadi, M., & Vahabzade, G., 2013. Spatial variability of heavy metals in surficial sediments: Tajan River Watershed, Iran. *Sustainability of Water Quality and Ecology*, 1–2, 48–58.

Gök, G. & Gürbüz, O.A., 2020. Application of geostatistics for grid and random sampling schemes for a grassland in Nigde, Turkey. *Environmental Monitoring and Assessment*, 192 (5), 300.

Grekousis, G., 2020. *Spatial Analysis Methods and Practice: Describe – Explore – Explain through GIS*. 1st edn. Cambridge University Press.

Hernández, Á., Arellano, E.C., Morales-Moraga, D., & Miranda, M.D., 2016. Understanding the effect of three decades of land use change on soil quality and biomass productivity in a Mediterranean landscape in Chile. *CATENA*, 140, 195–204.

Hoss, M., Behnke, G.D., Davis, A.S., Nafziger, E.D., & Villamil, M.B., 2018. Short Corn Rotations Do Not Improve Soil Quality, Compared with Corn Monocultures. *Agronomy Journal*, 110 (4), 1274–1288.

Hu, N., Xu, C., Geng, Z., Hu, F., Li, Q., Ma, R., & Wang, Q., 2022. The interplay of particle properties and solution chemistry on aggregation kinetics of soil nanoparticles. *Journal of Soils and Sediments*, 22 (6), 1761–1772.

İmamoğlu, M. & Sertel, E., 2016. Analysis of Different Interpolation Methods for Soil Moisture Mapping Using Field Measurements and Remotely Sensed Data. *International Journal of Environment and Geoinformatics*, 3 (3), 11–25.

Kim, H.N. & Park, J.H., 2024. Monitoring of soil EC for the prediction of soil nutrient regime under different soil water and organic matter contents. *Applied Biological Chemistry*, 67 (1), 1.
Kravchenko, A.N., 2003. Influence of Spatial Structure on Accuracy of Interpolation Methods. *Soil Science Society of America Journal*, 67 (5), 1564–1571.
Larney, F.J., Pearson, D.C., Blackshaw, R.E., & Lupwayi, N.Z., 2017. Soil surface cover on irrigated rotations for potato (*Solanum tuberosum* L.), dry bean (*Phaseolus vulgaris* L.), sugar beet (*Beta vulgaris* L.), and soft white spring wheat (*Triticum aestivum* L.) in southern Alberta. *Journal of Soil and Water Conservation*, 72 (6), 584–596.
Li, J. & Heap, A.D., 2014. Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling & Software*, 53, 173–189.
Liu, P., Zhao, H., Wang, L., Liu, Z., Wei, J., Wang, Y., Jiang, L., Dong, L., & Zhang, Y., 2011. Analysis of Heavy Metal Sources for Vegetable Soils from Shandong Province, China. *Agricultural Sciences in China*, 10 (1), 109–119.
Liu, S. & Shang, S., 2026. A combined spatial interpolation method of co-Kriging with inverse distance weighting and random forest for soil water and salt in arid oasis. *Journal of Hydrology*, 664, 134569.
Lu, A., Wang, J., Qin, X., Wang, K., Han, P., & Zhang, S., 2012. Multivariate and geostatistical analyses of the spatial distribution and origin of heavy metals in the agricultural soils in Shunyi, Beijing, China. *Science of The Total Environment*, 425, 66–74.
Mackman, T.J. & Allen, C.B., 2010. Investigation of an adaptive sampling method for data interpolation using radial basis functions. *International Journal for Numerical Methods in Engineering*, 83 (7), 915–938.
Miao, Y., Li, H., Xue, L., Shen, C., & Wang, F., 2026. Application of a semi-variogram-based KNN algorithm in the spatial prediction of soil heavy metals. *Environmental Pollution*, 390, 127436.
Nogiya, M., Moharana, P.C., Meena, R., Yadav, B., Jangir, A., Malav, L.C., Sharma, R.P., Kumar, S., Meena, R.S., Sharma, G.K., Jena, R.K., Mina, B.L., & Patil, N.G., 2024. Spatial variability of soil variables using geostatistical approaches in the hot arid region of India. *Environmental Earth Sciences*, 83 (14), 432.
Oliver, M.A. & Webster, R., 2014. A tutorial guide to geostatistics: Computing and modelling variograms and kriging. *CATENA*, 113, 56–69.
Oliver, M.A. & Webster, R., 2015. *Basic steps in geostatistics: the variogram and kriging*. Springer.
O'Sullivan, D. & Unwin, D.J., 2010. *Geographic Information Analysis*. 1st edn. Wiley.
Panday, D., Ojha, R.B., Chalise, D., Das, S., & Twanabasu, B., 2019. Spatial variability of soil properties under different land use in the Dang district of Nepal. *Cogent Food & Agriculture*, 5 (1), 1600460.
Quino Lima, I., Ramos Ramos, O., Ormachea Muñoz, M., Quintanilla Aguirre, J., Duwig, C., Maity, J.P., Sracek, O., & Bhattacharya, P., 2020. Spatial dependency of arsenic, antimony, boron and other trace elements in the shallow groundwater systems of the Lower Katari Basin, Bolivian Altiplano. *Science of The Total Environment*, 719, 137505.
Safaee, S., Libohova, Z., Kladivko, E.J., Brown, A., Winzeler, E., Read, Q., Rahmani, S., & Adhikari, K., 2024. Influence of sample size, model selection, and land use on prediction accuracy of soil properties. *Geoderma Regional*, 36, e00766.
Saito, H. & McKenna, S.A., 2007. Delineating high-density areas in spatial Poisson fields from strip-transect sampling using indicator geostatistics: Application to unexploded ordnance removal. *Journal of Environmental Management*, 84 (1), 71–82.
Schmidinger, J., Schröter, I., Bönecke, E., Gebbers, R., Ruehlmann, J., Kramer, E., Mulder, V.L., Heuvelink, G.B.M., & Vogel, S., 2024. Effect of training sample size, sampling design and prediction model on soil mapping with proximal sensing data for precision liming. *Precision Agriculture*, 25 (3), 1529–1555.
Sharma, D., Inbaraj, M.P., Naz, A., & Chowdhury, A., 2024. Fate, source apportionment and fractionation of potentially toxic elements in agricultural soil around a densely populated,

semiarid urban center of India: baseline study and ecological risk assessment. *Environmental Geochemistry and Health*, 46 (6), 207.
Shoumik, B.A.A., Błońska, E., & Lasota, J., 2025. Soil Quality Index According to Diverse Land Use Systems Across the Europe. *Land Degradation & Development*, 36 (5), 1467–1482.
Sun, Y., Dai, H., Moayedi, H., Nguyen Le, B., & Muhammad Adnan, R., 2024. Predicting steady-state biogas production from waste using advanced machine learning-metaheuristic approaches. *Fuel*, 355, 129493.
Szatmári, G., Pásztor, L., & Heuvelink, G.B.M., 2021. Estimating soil organic carbon stock change at multiple scales using machine learning and multivariate geostatistics. *Geoderma*, 403, 115356.
Tepanosyan, G., Sahakyan, L., Zhang, C., & Saghatelyan, A., 2019. The application of Local Moran's I to identify spatial clusters and hot spots of Pb, Mo and Ti in urban soils of Yerevan. *Applied Geochemistry*, 104, 116–123.
USEPA, 1996. Method 3540C, Soxhlet extraction. *Test Methods for Evaluating Solid Waste, Physical/Chemical Methods*, Revision 3, 1–8.
Watkins, K.B. & Lu, Y.-C., 1998. Economic and Environmental Tradeoffs Among Alternative Seed Potato Rotations. *Journal of Sustainable Agriculture*, 13 (1), 37–53.
Webster, R. & Oliver, M.A., 2007. *Geostatistics for Environmental Scientists*. 1st edn. Wiley.
Wei, J.-B., Xiao, D.-N., Zeng, H., & Fu, Y.-K., 2008. Spatial variability of soil properties in relation to land use and topography in a typical small watershed of the black soil region, northeastern China. *Environmental Geology*, 53 (8), 1663–1672.
Yang, X., Jia, C., Yao, Y., Yang, T., & Shao, S., 2024. Precise management and control around the landfill integrating artificial intelligence and groundwater pollution risks. *Chemosphere*, 364, 143185.
Yaşar Korkanç, S., 2018. Effects of the land use/cover on the surface runoff and soil loss in the Niğde-Akkaya Dam Watershed, Turkey. *CATENA*, 163, 233–243.
Zhang, C., Luo, L., Xu, W., & Ledwith, V., 2008. Use of local Moran's I and GIS to identify pollution hotspots of Pb in urban soils of Galway, Ireland. *Science of The Total Environment*, 398 (1–3), 212–221.
Zhang, D., Zhao, Y., Qi, H., Shan, L., Chen, G., & Ning, T., 2024. Effects of Micro-Topography and Vegetation on Soil Moisture on Fixed Sand Dunes in Tengger Desert, China. *Plants*, 13 (11), 1571.
Zhang, H., Lu, L., Liu, Y., & Liu, W., 2015. Spatial Sampling Strategies for the Effect of Interpolation Accuracy. *ISPRS International Journal of Geo-Information*, 4 (4), 2742–2768.
Zou, G., Xue, K., Huang, D., Su, C., & Sun, J., 2010. The comparison and study of small sample data spatial interpolation accuracy. *In*: *2010 Sixth International Conference on Natural Computation*. Presented at the 2010 Sixth International Conference on Natural Computation (ICNC), Yantai, China: IEEE, 1897–1900.