International Journal of Geography, Geology and Environment

P-ISSN: 2706-7483
E-ISSN: 2706-7491
IJGGE 2023; 5(2): 11-16
Received: 12-04-2023
Accepted: 16-05-2023

Sangita Singh
University School of Environment Management, Guru Gobind Singh Indraprastha University, New Delhi, India

Kiranmay Sarma
University School of Environment Management, Guru Gobind Singh Indraprastha University, New Delhi, India

Corresponding Author:
Sangita Singh
University School of Environment Management, Guru Gobind Singh Indraprastha University, New Delhi, India

Spatial variability of soil parameters: A geostatistical approach

Sangita Singh and Kiranmay Sarma

DOI: https://doi.org/10.22271/27067483.2023.v5.i2a.166

Abstract
Prediction models have been widely used to create a statistical model and understand the relationship between environmental variables and soil attributes. The present study was conducted to access the spatial variability in soil parameters by using geostatistical techniques. The study area selected was the National Capital Territory (NCT) of Delhi within which 22 sampling points were chosen for sample collection from the surface layer. The parameters (pH, Electrical Conductivity, Soil Moisture (SM)%, and Soil Organic Carbon (SOC) % were analyzed against the prediction estimates, and a regression analysis supports the inter-relation of the observed and predicted set of values. The method of interpolation used in the study was RBF (Radial Basis Function) which was carried out by using the ArcGIS software. The cross-validation of the data set was also analyzed by calculating the Mean and Root Mean Square Error (RMSE). According to the results, sample distance is adequate for interpolation and the RBF can clearly show the geographical distribution of soil attributes.

Keywords: Spatial variability, soil, geostatistical, radial basis function, Delhi

1. Introduction
Reliable data on the spatial distribution of soil qualities influencing both landscape processes and services are necessary for sustainable land management. The majority of soil maps in India were created using traditional methodologies, and little research has been done thus far to utilize modern predictive spatial tools. It has become more crucial to spatially map soil characteristics at unobserved places using statistical inference. The advent of geostatistics, which allowed scientists to precisely interpolate [1] spatial patterns of soil attributes, contributed to the usefulness of such maps. In this, geographic and non-spatial soil inference systems are combined with field and laboratory observational methods to create and populate spatial soil information systems. The spatial distribution [4] of soil properties is determined through field sampling, and the soil properties in regions that were not tested are then estimated using surface grid point data that have been interpolated. Contrarily, current traditional soil survey procedures are lengthy and expensive, and the resulting soil databases aren't accurate or comprehensive enough to support extensive and reliable use of soil information in spatial data. The variability of soil productivity and soil health is significantly influenced by soil nutrients, one of the major factors influencing soil quality. Employing an interpolation technique, the distribution of soil parameters can be plotted. The words "deterministic" and "stochastic" have been used to identify a variety of spatial interpolation [5] techniques that have been created. While stochastic interpolation techniques like kriging, RBF, and Inverse Distance Weighing (IDW) approaches do provide assessments of prediction error, deterministic interpolation techniques like Thiessen, density estimation, inverse-distance-weighted, and splines do not. In interpolation techniques, the spatial statistical technique evaluates the autocorrelation that is frequently seen in geographic data, where data values from nearby locations are more similar than data values from distant locations. In soil science, stochastic [7] methods produce reliable results. The RBF method produces results that are more accurate than those produced by kriging and IDW interpolation. In this article, RBF is precisely used to create a spatial prediction of the soil parameters.
RBF [3, 17] is a geostatic approach to interpolation that has been shown to be sufficiently reliable for extrapolating values from sampled data to non-sampled places. It offers the most accurate linear unbiased estimates, details the distribution of the estimation error, and demonstrates significant statistical benefits.

If a particular set of soil samples adequately represents the research area, using the geostatistical interpolation technique also lowers the expenses of field collection and laboratory analysis. However, sufficient sampling data and precise geographical interpolation are necessary for the validity of spatial variability maps. RBF algorithms may project values that can range above or below the measured values' maximum or lowest. The approach's estimated values are based on a mathematical formula that reduces the overall curvature of the surface, producing flat surfaces. To employ spatial statistical approaches to analyze the soil characteristics and geographical distribution [11,16] of soil parameters at the surface layer, a pilot study was carried out. The objective of the study is to assess the possibilities for employing geostatistical techniques to measure soil characteristics and their geographical variability in the NCT of Delhi. Knowledge of the geographical variability of soil nutrients is crucial for the long-term management of soil fertility.

2. Methodology

2.1 Study area: With a 1483 km2 area and an average elevation of 213-290 m above mean sea level, Delhi is located in a latitudinal extent of 28°23'17"-28°53'00" N and longitudinal extent of 76°50'24"-77°20'37" E (Figure 1). New Delhi, Central Delhi, West Delhi, North Delhi, North West Delhi, South Delhi, South East Delhi, East Delhi, North East Delhi, and Shahdara are the eleven districts that collectively constitute the area. Gurgaon, Noida, Faridabad, and Ghaziabad are a few of the nearby satellite cities that finally sprang up around Delhi. Uttar Pradesh borders it on the east, while Haryana borders it on the north, west, and south. From a larger geographical perspective, the research region is situated between the Himalayas in the north, the Aravalli mountains in the south, and the Yamuna River in the east. East Delhi and West Delhi are the two areas of the city that are separated by the Yamuna River. The ridge, which is a continuation of the Aravalli Range, abuts Delhi on its southern boundary and rises to the city's center. The majority of the vegetation on the ridge is thorny. Due to the aforementioned geographic expanse, there are dry, chilly winters (1-3 °C) and summers with extremely high peak temperatures (45-47 °C). The primary climatic component on which most of the agriculture depends, with an average rainfall of 790 mm, is the monsoon season. According to the 2011 Census of India, Delhi is home to 16.75 million people. Delhi has a population density of approximately 11,297 people per km^2. Delhi's total area covered by trees and forests (176.2 km^2) is around 296.2 km^2 or roughly 20% of the entire area. The predominant vegetation is thorny scrub, which is indicative of semi-arid conditions. Alkaline and saline soils make up the majority of the prevalent soil type [18].

2.2 Soil sampling: A stratified random sampling procedure was followed for the sample collection (0-10 cm) during April 2018. A total of 22 sites were selected and spotted with the help of a hand-held GPS E-Trex20. Five soil cores were used to collect the undisturbed soil samples and they were thoroughly combined to create a composite soil sample. For the laboratory analysis samples were air-dried and put through a 2 mm sieve. The parameters selected for the study include soil pH, soil EC, OM%, and SM%.

2.3 Integration of field and GIS platform: Deterministic interpolation methods were used in the ArcGIS software's Geostatistical Analyst [10, 20] to create thematic maps of soil parameters. The spatial distribution was produced using the deterministic interpolation approach with a degree of smoothing (radial basis functions, RBF). Projected values may vary above or below the maximum or lowest of the measured sample values when using the RBF method since the surface passes across each measured sample value.

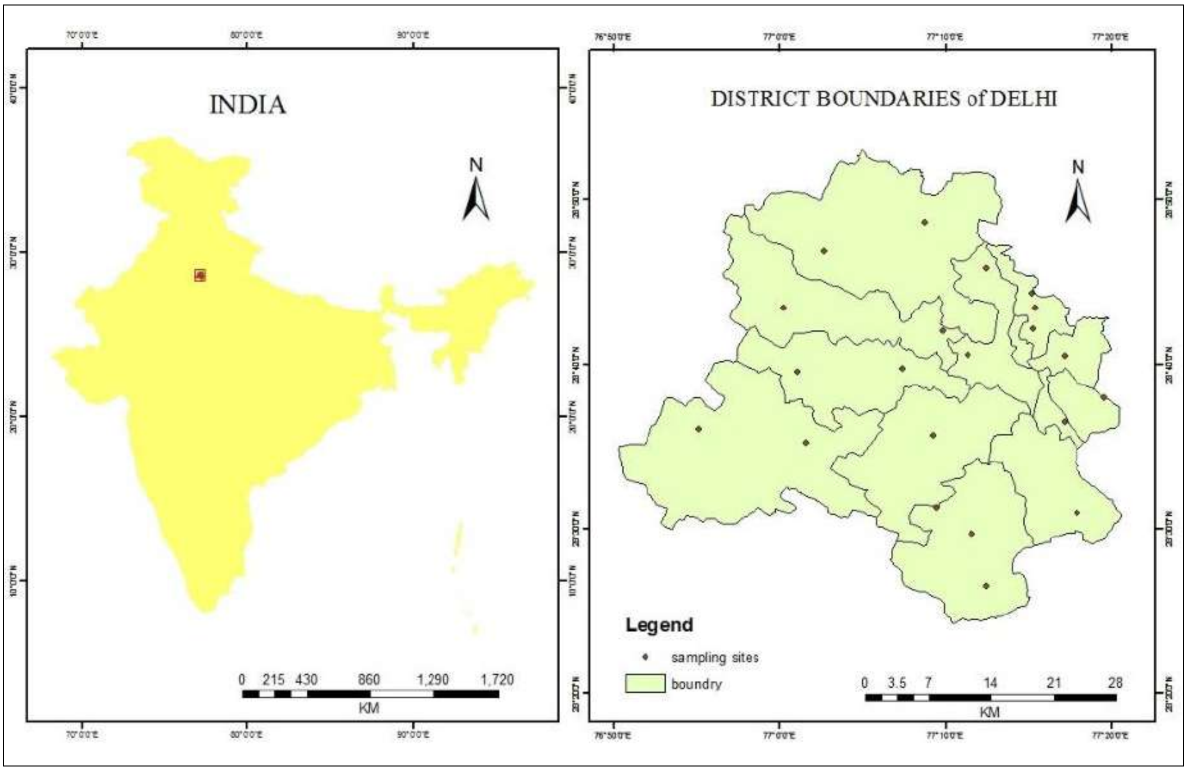

Fig 1: Study area location with sampling points

The soil parameters in unsampled places are predicted using the RBF model. RBF is used to interpolate data points in a group of multi-dimensional space, and prior research suggests [3] that it is a good method for interpolating soil properties. RBF minimizes the surface's overall curvature while fitting a surface through the measured sample values. When there is a sharp change in the surface values within proximity, RBF is ineffectual. In this study, the most popular RBF, CRS (completely regularised spline), was chosen.

A radial basis function (RBF) is a mathematical function that takes a vector as input and outputs a scalar value. The name "radial" refers to the fact that the function's value depends on the distance between the input vector and a center point. The general form of an RBF is:

$$\phi(r) = f(\|x - C\|)$$

Here, ϕ(r) represents the output value of the RBF, ||x - c|| denotes the Euclidean distance between the input vector x and a center point c, and f is a univariate function that determines the shape of the RBF.

2.4 Data validation: RBF [3] interpolation method performance was assessed using the cross-validation approach. Two datasets made up of the sample points were randomly created; one was used to train a model, and the other was used to evaluate the model. The training and validation sets must overlap in successive rounds so that each data point can be validated to reduce variability [12,13]. For each of the generated topsoil prediction [9] maps, the root-mean-squared error (RMSE) and the goodness of prediction were determined as indicators of accuracy and performance, respectively. When its value tends to zero, the least RMSE [20] values are used to assess a prediction model's fitness.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(O_i - S_i)^2}{N}}$$

Where 0i is the observed value and Si is the predicted value, N is the Number of samples.

3. Results and Discussion

The statistical [14] methods were analyzed by initially visualizing the histograms of the soil samples in the GIS environment (for normal distribution) and calculating the descriptive statistics (Table 1).

Table 1: Descriptive statics of the soil parameters

Variable	Min	Max	Mean	S.D.	Skew.	Kurto.	1st Quartile	Median	3rd Quartile
pH	7.3	8.9	8.29	0.34	-0.97	4.33	8.2	8.3	8.5
EC	0.001	0.07	0.013	0.01	2.24	7.32	0.003	0.006	0.017
SM%	0.11	18.87	2.16	4.08	3.37	14.04	0.45	0.85	1.58
SOC%	0.41	5.29	2.22	1.27	0.87	2.98	1.18	1.88	3

3.1 pH

Measured using an electrometric technique that involves comparing the test solution's e.m.f. (millivolts) to that of the reference buffer. Generally referred to as the soil solution's active hydrogen ion's (H+) negative logarithm. It gauges the sodicity, acidity, or neutrality of the soil. It is significant because it improves the soil's nutritional status. Acidic nature is generally indicated by low pH readings, and alkaline nature [19] by higher values. In the study, it was found that the majority of the sites (Fig. 3) indicated that the soil was alkaline [18]. Values varied from 7.3 to a maximum of 8.9, with 8.29 serving as the average (Table 1). According to another source [8] the hydrogen and aluminium ions that remain after basic cations and other nutrients are lost to erosion and leaching are what cause the soil to become acidic. A particular pH range favors the majority of microbial activities and soil processes.

3.2 EC

It is measured to determine the ionic content of the soil sample. The value, which is often expressed in dS.m-1, provides details [19] on the total quantity of the soluble salts. Salted soils are often divided into groups based on two criteria: total soluble salt content and sodium absorption ratio. Analyses can be used to determine very high values in salts with increased organic matter or sodium contents. From the result obtained the range was observed as 0.001-0.07 dS.m^{-1} with 0.01 as deviation from the mean range. The variations are mainly due to the salinity level, water content, cation exchange capacity, and temperature. The trend of conductivity is quite different from soil pH as it is more variable in the study area.

3.3 SM%

Soil moisture is the water content present in soil particles which is also an important parameter. The exceeding and lower values both hinder the growth phase [19] of soil microbes as well as plants. The values range was observed to be 0.11-18.76% (mean=2.16) with a CV of 16.713. According to the results observed, it can be concluded that the variations were in the moderate range throughout the study area. Temperature variations and vegetation cover can affect a lot to this parameter [19]. By regulating water retention and flow in the soil profile, soil type can affect soil moisture. For instance, soils with a high percentage of clay tend to have a higher ability to hold onto water, whereas sandy soils may have a higher rate of water movement but have a lesser ability to hold onto water.

3.4 SOC%

Gives the amount of organic [2] carbon percent present in the soil, calculated by the Walkley and Black [6] method. It is one of the important parameters [19] which decides the health of the soil. The higher values correspond [2] to a good humic and porous soil that is required for plant growth. It also facilitates soil microbial growth. The values of SOC% are quite variable likewise in EC, here the results show different patches which are quite separable and are not have an in-between value. The pattern of only this variable was quite different from the rest, most of the area has shown to have lower values at maximum locations to moderate ranges at few. The range of SOC% was observed as 0.32-5.29, with a mean of 2.22 and 1.27 as standard deviation from the observations. Temperature is an important phenomenon as Warmer temperatures tend to boost microbial activity, which speeds up the decomposition of organic matter and lowers SOC levels. Cooler temperatures, on the other hand, can sluggish microbial activity and decrease the rate of decomposition. SOC levels can also be impacted by soil moisture levels. The microbial activity in soils has a tendency to be lower when they are overly dry. On the other hand, excessive soil moisture can impede decomposition since there is less oxygen available, which raises SOC levels. SOC levels can be impacted by the presence and development of vegetation since plants can enrich the soil with organic matter through their roots and above-ground biomass. Plants may release more organic compounds into the soil through their root systems during the growing season when they are actively absorbing nutrients, which can raise SOC levels. SOC levels can shift as a result of actions like tillage or changes in land usage. In general, soil disturbance can speed up the breakdown of organic matter and result in lower SOC levels. For a better understanding of the measured and predicted value, a regression function (Table 2) is applied and the errors were predicted from the same (Fig. 2).

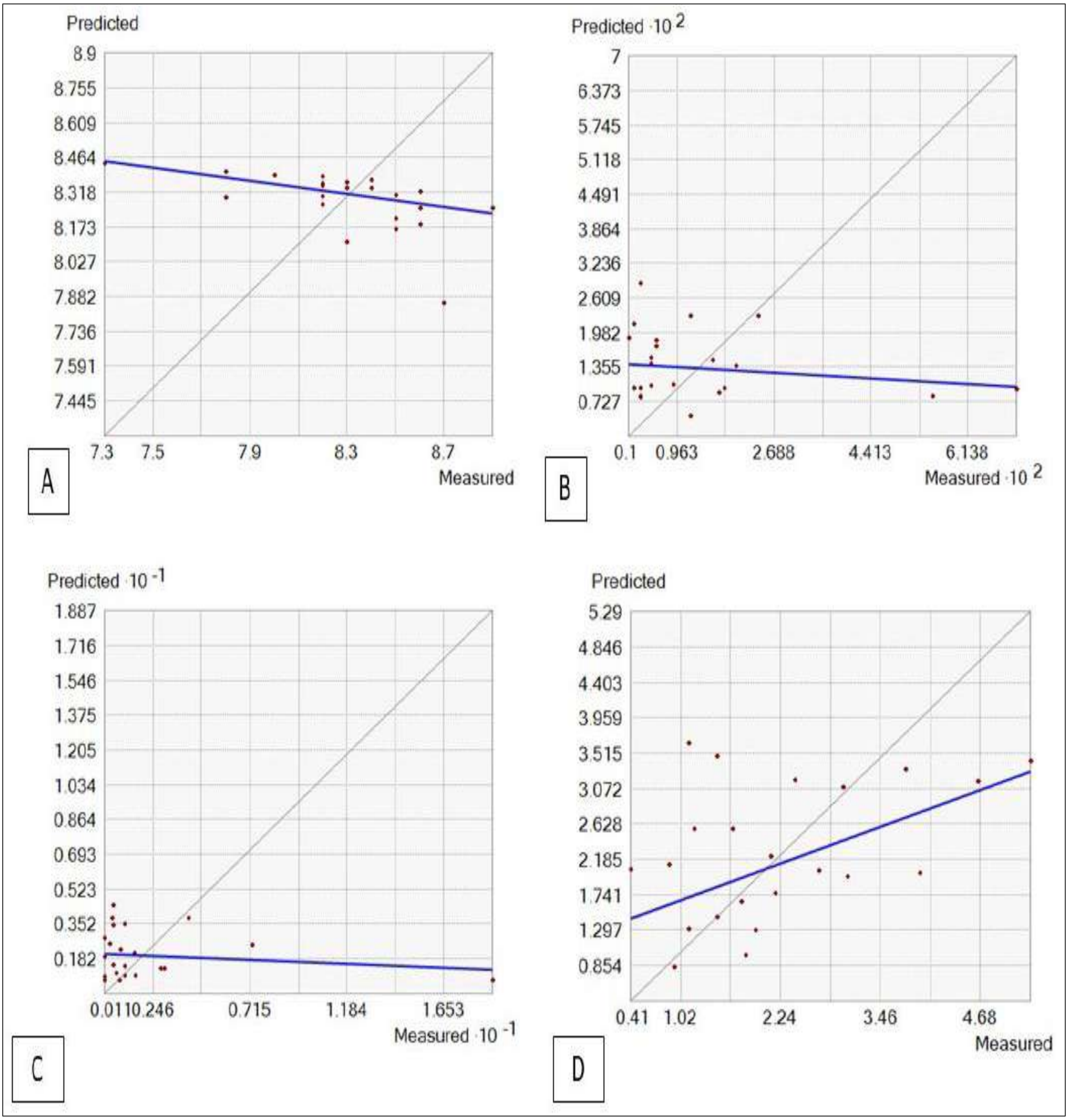

Fig 2: Scatterplot of measured values with the predicted values

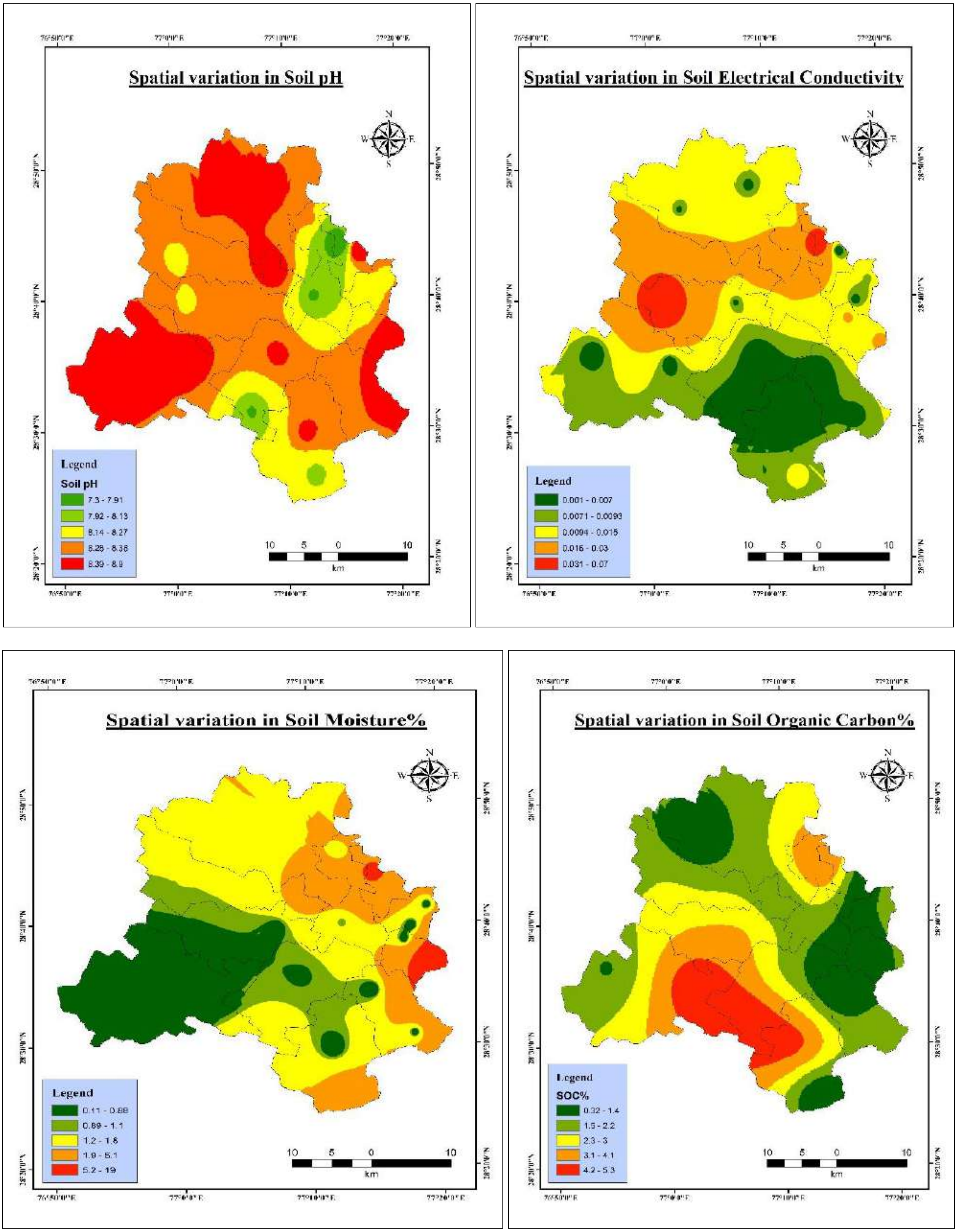

Fig 3: Geostatistical variation in soil parameters with Radial Basis Function.

The regression function applied provides a better idea of the actual ground variations [11, 15] and their relationship with applying distance. According to the results obtained the best error output was of soil EC with a mean of 0.00051 and RMSE of about 0.0192. The results were finally analyzed on the basis of RMSE error which gives a trend of EC>pH>SOC%>SM%. In Table 2, Y represents the dependent variable and X is the explanatory variable.

Table 2: Interpolation parameters with Kernel function (Completely Regularized Spline)

Variable	Mean error	RMSE	Regression function
pH	-0.010	0.422	Y= -0.136 * x + 9.444
EC	0.00051	0.0192	Y= -0.0595 * x + 0.0141
SM%	-0.0817	4.350	Y=-0.0414 * x + 2.0642
SOC%	0.0664	1.179	Y= 0.3776 * x + 1.2865

4. Conclusion

From the aforementioned findings, it can be inferred that the geostatistical technique, which applies both descriptive statistics and predictive analysis, enhanced the field-scale description of spatial variability. The distribution of soil properties in non-sampled sites based on sampled data was successfully explained by the RBF maps of soil properties. Precision information, environmental monitoring, and model development all depend on having a thorough understanding of the spatial distribution and precise mapping of soil attributes at a broad scale. This work demonstrated how soil parameters, including pH, electrical conductivity (EC), soil moisture (SM %), and soil organic carbon (SOC), were fitted into geostatistical models. It was demonstrated through cross-validation using RMSE that spatial prediction of soil parameters is preferable to assuming the mean. The findings imply that the sample distance in this study is enough for interpolation and that the RBF interpolation can directly disclose the geographical distribution of soil attributes. Future research is nonetheless required to define spatial variability on a broader scale and comprehend the variables influencing the spatial variability of soil parameters. Since a huge sample size is practically unachievable, interpolation techniques are an excellent way to forecast the sites that were not sampled. For the soil variable prediction map, the RBF interpolation method's results were satisfactory. Due to these findings, geostatistical analysis is suggested as a method for future soil sampling to explore the spatial variability of soil variables.

5. Acknowledgments

The authors would like to pay their gratitude towards UGC for providing financial assistance to the first author and are also thankful to the University School of Environment Management, Guru Gobind Singh Indraprastha University, New Delhi, India, for providing the laboratory and other necessary facilities.

6. References

1. Antwi M, Duker AA, Fosu M, Abaidoo RC. Geospatial approach to study the spatial distribution of major soil nutrients in the Northern region of Ghana. Cogent Geoscience. 2016;2(1).
2. Mondal A, Khare D, Kundu S, Mondal S, Mukherjee S, Mukhopadhyay A. Spatial soil organic carbon (SOC) prediction by regression kriging using remote sensing data. The Egyptian Journal of Remote Sensing and Space Science. 2017;20(1):61-70.
3. Bhunia GS, Shit PK, Maiti R. Spatial variability of soil organic carbon under different land use using radial basis function (RBF). Modeling Earth Systems and Environment. 2016;2:1-8.
4. Myers DE. Spatial interpolation: an overview. Geoderma. 1994;62(1-3):17-28.
5. Lopez GF, Jurado EM, Pena BJM, Garcia TL. Using geostatistical and remote sensing approaches for mapping soil properties. European Journal of Agronomy. 2005;23 (3):279-289.
6. Walkley A, Black IA. An examination of the Degtjareff method for determining organic carbon in soils: Effect of variations in digestion conditions and of inorganic soil constituents. Soil Science. 1934;63:251-263.
7. Yasrebi J, Saffari M, Fathi H, Karimian N, Moazallahi M, Gazni R. Evaluation and comparison of ordinary kriging and inverse distance weighting methods for prediction of spatial variability of some soil chemical parameters. Research Journal of Biological Sciences. 2009;4(1):93-102.
8. Tziachris P, Metaxa E, Papadopoulos F, Papadopoulou M. Spatial Modelling and Prediction Assessment of Soil Iron Using Kriging Interpolation with pH as Auxiliary Information. ISPRS International Journal of Geo-Information. 2017;6(9):283.
9. Orman EE. Improving the prediction accuracy of soil mapping through geostatistics. International Journal of Geosciences. 2012;3:574-590.
10. Johnston K, Ver H, Jay M, Krivoruchko K, Lucas N. Using ArcGIS Geostatistical Analyst. Environmental Systems Research Institute, Redlands, CA; c2001.
11. Vasu D, Singh SK, Sahu N, Tiwary P, Chandran P, Duraisami VP, *et al*. Assessment of spatial variability of soil properties using geospatial techniques for farm level nutrient management. Soil and Tillage Research. 2017;169:25-34.
12. Tripathi R, Nayak AK, Shahid M, Raja R, Panda BB, Mohanty S, *et al*. Characterizing spatial variability of soil properties in salt affected coastal India using geostatistics and kriging. Arabian Journal of Geosciences. 2015;8(12):10693-10703.
13. Dongare VT, Reddy GP, Kharche VK, Ramteke IK. Spatial variability of soil nutrients under sugarcane cropping system in semi-arid tropics of western India using geostatistics and GIS. Journal of Soil and Water Conservation. 2022;21(1):67-75.
14. Eljebri S, Mounir M, Faroukh AT, Zouahri A, Tellal R. Application of geostatistical methods for the spatial distribution of soils in the irrigated plain of Doukkala, Morocco. Modeling earth systems and environment. 2019;5:669-687.
15. Gangopadhyay SK, Reddy GP. Spatial variability of soil nutrients under the rice-fallow system of eastern India using geostatistics and Geographic Information System. Journal of Soil and Water Conservation. 2022;21(1):55-66.
16. Khan MZ, Islam MR, Salam ABA, Ray T. 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; c2021. p. 1-19.
17. Sarmadian F, Ali K, Antonia O, Ghavamoddin Z, Hossein J. Mapping of spatial variability of soil organic carbon based on radial basis functions method. ProEnvironment. 2014;7:3-9.
18. Singh S, Sarma K. Mapping Surface Soil Characteristics of Barren Land by Using Geospatial Technology in NCT of Delhi. Environment & We an International Journal of Science & Technology. 2020;15(1):15-27.
19. Panda SC. Agronomy. Agrobios, India; c20100. p. 99-137.
20. ESRI. ArcGIS^{TM} Geostatistical Analyst: Statistical Tools for Data Exploration, Modeling and Advanced Surface Generation; c2001. http://www.esri.com