Original Research Article

Study of Advanced Techniques to Predict the Soil Properties

ABSTRACT

Information about soil properties helps the farmers to adopt e□ective and e□cient farming practices, which can increase higher yields with optimum usage of farm resources. An attempt has been made in this paper to predict soil properties using geospatial kriging approaches. This study mainly focuses on predicting soil pH using different kriging methods. Soil pH dramatically affects many other soil processes, such as nitrification and denitrification, mineralization, precipitation, and dissolution of soil organic matter. Total of seven kriging semivariogram models, namely spherical, circular, exponential, Gaussian, and linear, while two models of universal kriging, such as linear with linear drift and linear with quadratic drift, have been taken to interpolate the spatial soil pH. The performances of these entire models have been validated using mean error, and root mean square error. Spatial analysis revealed that Universal kriging outperformed ordinary kriging with less mean error and root mean square error, 0.016 and 0.52, respectively. The spatial analysis of soil mapping can be instrumental in adopting real-time and on-the-go soil precision practices.

Keywords: Kriging; soil pH; soil properties; spatial analysis

INTRODUCTION

It is essential to know the soil properties to apply the correct dose of soil nutrients to the soil in precision farming [8]. As it is known, the soil is a complex mixture of six natural elements, including inorganic, organic matter, soil biota, moisture content, soil buffer, and soil air. Insufficient knowledge about soil properties is one of the primary impediments to cultivating the field's crops. In addition, soil properties vary spatially, and the procedure for determining the soil properties in the lab is tedious. Therefore, there is a need to be studied and predicted to prepare soil maps that can be used to apply the current agronomical practices under the precision farming concept. The use of machine learning methods in various agricultural research areas has received more attention in recent years, and soil properties prediction is one of the most applied areas of these methods [1]. Soil "power of hydrogen", often referred to as soil reaction (pH) by scientists, is the most famous indicator of soil quality [6, 7]. A technique known as kriging has proven to be decisive for predicting values at an unknown location based on data collected from samples in a given area. Compared to other methods, this method provided prime linear unbiased approximation and information on the approximation error distribution and showed considerable statistical advantages [5].

Consequently, in this study, the unknown properties of the soil were interpolated using the kriging method. A spatial distribution map of different soil pH values was generated based on the predicted soil properties using the ordinary and universal kriging methods, which are the geostatistical analyst algorithm. Using the OK and UK methods for soil property interpolation and spatial mapping, they conducted a study that evaluated the optimum environmental variables to predict SOC, soil pH, and TN [2]. One of the researchers investigated 3D-DSM to assess clay content, volumetric water content, and soil organic matter using multiple proximal soil sensing techniques combined with a 3D regression kriging method [9]. The soil chemical properties of agricultural fields were determined through digital processes, which vary depending on the seasonal rainfall and the management practices [5]. Numerous digital map layers have been created based on the soil's chemical properties and the data collected in ArcGIS using a geostatistical tool, kriging, to predict the unknown measured values. Even though many advanced technologies, such as machine learning, deep learning, and remote sensing, have been successfully used by researchers to know the current state of soil types based on soil pH, the problem remains. Until recently, no investigation has found the use of soil mapping to signify soil reactions in the context of soil pH on a spatial scale. So, this study aimed to predict the soil pH using the different kriging methods using ArcGIS.

MATERIAL AND METHODS

Study area and sample collection

Due to the ease of collecting the soil samples, the study area has been selected as the experimental field of ICAR-CIAE, Bhopal. The experimental area was located in the same field for soil sample collection. The experimental area is black cotton soil with 12.6% sand, 32.7% silt, and 54.7% clay content. The soil samples were collected by the manual soil sampler at 0-15 cm depth, thoroughly mixed, stored, sun-dried, and pre-processed to determine the soil pH. The shape file of the study area has been created in ArcGIS software, shown in Fig. 1. Soil pH was determined using HANNAHI98195 portable meter as per standard.

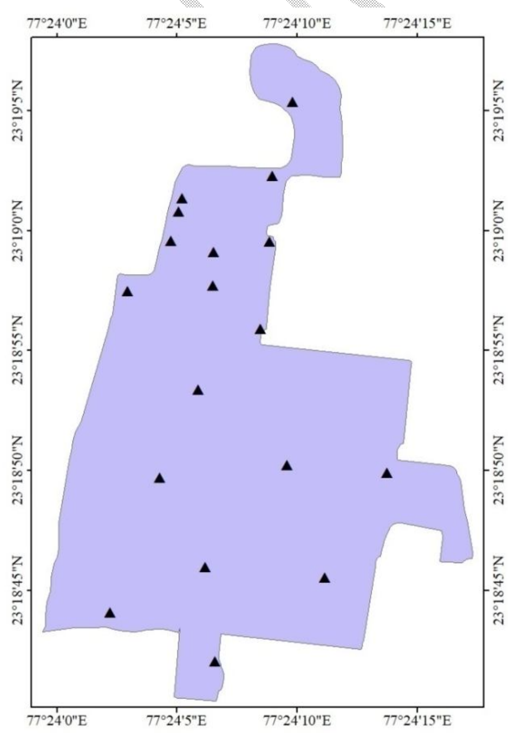


Fig. 1 Study area shape file in ArcMap

Spatial analysis

Kriging is a non-deterministic model that requires the examination of the underlying spatial behaviour of the point location data values. Kriging, in turn, informs how the interpolated raster surface will be estimated from the point data. Machine learning (ML) techniques frequently map soil characteristics using spatial variables and observed values from known locations. As a result, various approaches and geostatistical techniques have been developed to classify and map soil properties based on numerous methodologies. In interpolation and estimation, kriging models often correlate highly between linearity and mapped soil characteristics and inputs [4].

Ordinary kriging (OK)

There are a variety of kriging techniques, but OK is one of the most popular. The spatial prediction at the unknown location x_0 is given by $Z(x_0)$, which is a sum of the weightage value of the known measured values $Z(x_0)$. Gia Pham T. et al., 2019 have described OK as elegant and simple by the following equation (i).

$$Z(x_0) = \sum_{k=1}^{n} \lambda_k Z(x_k) \text{............ (i)}$$

where $Z(x_0)$ is the predicted value at an unknown location x_0, $Z(x_k)$ is the observed value at a known location x_k, λ_k is the weightage factor from a known location to an unknown location, and n is the quantity of the nearest neighbour. It is necessary to fit a model by the input data distribution to show the spatial relationship between neighbourhood points and the spatial continuation of the data.

Universal kriging (UK)

UK is a spatial interpolation methodology that united the regression of response factor variables on explanatory variables and the prediction residuals (Gia Pham T. et al., 2019). The following equation (ii) can be used to describe it.

$$Z(x_0) = \sum_{i=0}^{p} \hat{\beta}_i * q_i(x_0) + \sum_{k=1}^{n} \lambda_k e(x_k) \text{..........(ii)}$$

where $Z(x_0)$ is the predicted value at a non-known location x_0, $\hat{\beta}_i$ is the calculated deterministic model factor, λ_k is the weightage factor from a known location to an unknown location, $e(x_k)$ is the error at the known point x_k. This study applied the OK and UK methods using ArcGIS software.

Cross Validation

In this study, we used a cross-validation approach to test the accuracy of the models as per the values of root mean square error (RMSE) and mean error (ME) expressed in Eq. (iv) and Eq. (iii), respectively, to compare seven semivariogram models.

$$ME = \frac{1}{n} \sum_{i=1}^{n} (Z_{pi} - Z_{mi}) \text{ .... (iii)} \qquad ME = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Z_{pi} - Z_{mi})^2} \text{.......... (iv)}$$

Where, Z_{mi} is the measured value at the i^{th} position, Z_{pi} is the forecasted value at the i^{th} location and n is the number of samples taken.

RESULT AND DISCUSSION

In this study, we took seven semivariogram models to understand the comparative performance of kriging models to predict the soil pH spatially. A total of five models of ordinary kriging were selected, namely spherical, circular, exponential, Gaussian, and linear, while two models of universal kriging, such as linear with linear drift and linear with quadratic drift [3]. Soil pH prediction and spatial distribution maps were created using all seven models shown in Fig. 2. The spatial mean values of soil pH of the study area were found to be 7.16 ± 0.14, 7.17 ± 0.16, 7.17 ± 0.17, 7.16 ± 0.17, 7.17 ± 0.17 for spherical, circular, exponential, Gaussian, linear, linear with linear drift, and linear with quadratic drift respectively. The minimum and maximum spatial pH was interpolated to be 6.23 and 8.49 by universal linear with quadratic drift due to its nature as a second-order polynomial, which might be due to under-fitting or over-fitting data. The spherical model predicted mean spatial soil pH accurately out of other ordinary kriging models, while linear with linear drift predicted accurately in universal kriging. OK, and UK models were also cross-validated and compared with mean error, and root mean square error.

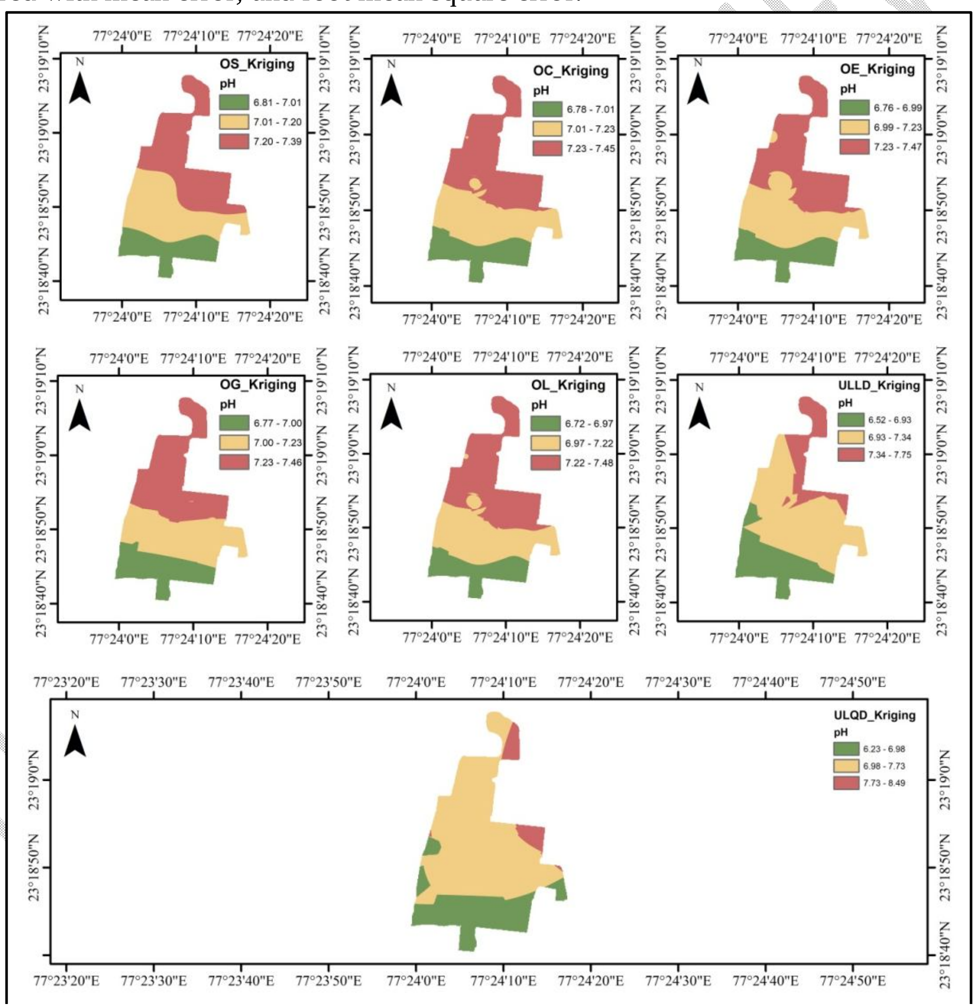
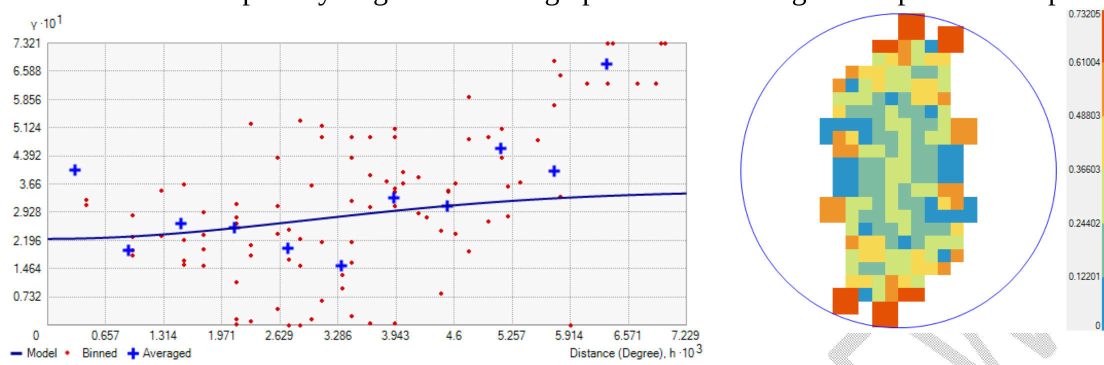


Fig. 2 Predicted maps of soil pH using semivariogram models

(OS: Ordinary spherical; OC: Ordinary circular; OE: Ordinary exponential; OG: Ordinary Gaussian; OL: Ordinary linear; ULLD: Universal linear with linear drift; ULQD: Universal linear with quadratic drift)

The semivariogram graph(shown in Fig. 3a) plotted the distance between the sample locations on the x-axis vs the semivariance on the y-axis. It represents the sample points closer to the model fitting line as more autocorrelated spatially, while farther apart points are less autocorrelated spatially. Fig. 3b shows a graph of a semivariogram map of the soil pH.



a b

Fig. 3 Experimental semivariogram graph and map

Table 1 shows below the geostatistical results summary of predicted pH values using different semivariogram models for ordinary and universal kriging. Universal kriging outperformed ordinary kriging with ME, and RMSE was 0.016 and 0.52, respectively (Table 2), due to which the UK uses trend surface to determine the maximum degree of the polynomial. Still, it does not use the coefficients from the trend surface.

Table 1 Geostatistical results summary

Ordinary kriging		**Universal kriging**	
Semivariogram model	**pH (mean ± sd)**	**Semivariogram model**	**pH (mean ± sd)**
Spherical	7.16 ± 0.14	Linear with linear drift	7.12 ± 0.25
Circular	7.17 ± 0.16	Linear with quadratic drift	7.17 ± 0.33
Exponential	7.17 ± 0.17		
Gaussian	7.16 ± 0.17		
Linear	7.17 ± 0.17		

Table 2 Prediction error summary

Prediction errors	**Ordinary kriging**	**Universal kriging**
Mean error (ME)	0.024	0.016
Root mean square error (RMSE)	0.59	0.52

From the analysis of different kriging models and the predicted results of the above study can be classified soil as slightly acidic to mildly alkaline, based on the mean spatial soil pH values, which is normal soil for crop cultivation without any reclamation practices. Geostatistical interpolation slightly signified a spatial variability of soil pH in the experimental study area.

CONCLUSION

In this study, we successfully compared the different kriging semivariogram model's performance, which can integrate the diverse information of geo-location mapping. Universal kriging predicted the best spatial soil pH with the minimum root mean square error than other semivariogram models. The spatial analysis of soil mapping can be instrumental in adopting real-time and on-the-go soil precision practices. This study can also be applied to other chemical properties to predict the spatial soil properties map.

REFERENCE

1. Ebid AM. 35 Years of (AI) in geotechnical engineering: state of the art. Geotechnical and Geological Engineering. 2021 Feb; 39(2):637-90.
2. Gia Pham T, Kappas M, Van Huynh C, Hoang Khanh Nguyen L. Application of ordinary kriging and regression kriging method for soil properties mapping in hilly region of Central Vietnam. ISPRS International Journal of Geo-Information. 2019 Mar 19;8(3):147.
3. Jaemyeong K, Yunsoo C, Hasu Y, Wonjong L. Optimal Space Interpolation Method for Continuous Marine Vertical Datum Based on WGS-84 Ellipsoid. Sensors and Materials. 2019 Jan 1;31(11):3917-30.
4. Kılıç M, Gündoğan R, Günal H, Cemek B. Accuracy Assessment of Kriging, artificial neural network, and a hybrid approach integrating spatial and terrain data in estimating and mapping of soil organic carbon. Plos one. 2022 May 26;17(5):e0268658.
5. Panday D, Maharjan B, Chalise D, Shrestha RK, Twanabasu B. Digital soil mapping in the Bara district of Nepal using kriging tool in ArcGIS. PloS one. 2018 Oct 26;13(10):e0206350.
6. Rezaei SA, Gilkes RJ, Andrews SS. A minimum data set for assessing soil quality in rangelands. Geoderma. 2006 Dec 1;136(1-2):229-34.
7. Volchko Y, Norrman J, Rosèn L, Norberg T. A minimum data set for evaluating the ecological soil functions in remediation projects. Journal of soils and sediments. 2014 Nov;14:1850-60.
8. Shit PK, Bhunia GS, Maiti R. Spatial analysis of soil properties using GIS based geostatistics models. Modeling Earth Systems and Environment. 2016 Jun;2:1-6.
9. Zhang Y, Ji W, Saurette DD, Easher TH, Li H, Shi Z, Adamchuk VI, Biswas A. Three-dimensional digital soil mapping of multiple soil properties at a field-scale using regression kriging. Geoderma. 2020 May 1;366:114253.