Agropedology 2022, 32 (01), 122-133
doi.org/10.47114/j.agroped.2022.jun9

Use of Soil Health Card Data for Nutrient Mapping: A Case Study of Bemetara District, Chhattisgarh

Adyasha Priyadarshini[1], R.K. Naitam[2], Nirmal Kumar[2*], K. Tedia[1], V.N. Mishra[1], Annu Singh[1], G.K. Srivastava[1] and R.R. Saxena[1]

[1] College of Agriculture, Indira Gandhi Krishi Vishwavidyalaya, Raipur-492 012
[2] ICAR-National Bureau of Soil Survey and Land Use Planning, Nagpur-440 033

Abstract: The soil health card programme has been undertaken by all Indian states for sustainable crop production. High resolution nutrient maps prepared from the measured point data will be of great help for fertilizer management. The present study used the popular method of interpolation *i.e.,* ordinary kriging (OK) in mapping major soil nutrients of Bemetara district from national soil health card scheme data. A total of 14491 geo-referenced soil health card data were used for the study. The soils were neutral in reaction (pH ranging 6.5 to 7.9 with mean of 7.2) and non-saline. The soils were mostly low to medium in organic carbon (mean 0.69 %) and available phosphorous (mean 17.24 kg ha^{-1}), low in available nitrogen (mean 228 kg ha^{-1}) and medium to high in available potassium (mean 390 kg ha^{-1}). Semivariogram modeling was done with 70 per cent of samples and the output maps were validated with the rest of 30 per cent samples. The semivariogram generated for the soil nutrients during interpolation approach (OK) showed poor spatial dependency among the points resulting in lower accuracies in nutrient mapping.

Key words: *Soil health card, ordinary kriging, interpolation, semivariogram*

Introduction

Balanced fertilizer use is the key to enhance use efficiency of the plant nutrients for maintaining the soil productivity. It aims at application of fertilizers in optimum quantities and in right proportion through appropriate methods, which results in sustenance of soil fertility and crop productivity. Due to insufficient and unbalanced use of plant nutrients by fertilizers and manures, the soil fertility status of both agriculturally advanced irrigated regions and less endowed rainfed regions in India has been depleted. There will be a greater drain on native soil fertility unless nutrients are supplied judiciously, and the soil will be unable to maintain high crop production in the years to come unless nutrients are supplied judiciously. Fertilizer use in India is highly skewed towards nitrogen. In 2012–13, the ratio of NPK use in India reached 8.2:3.2:1, which is more imbalanced compared to the early 1970s when the ratio was 6:1.9:1.

Site-specific nutrient management allows farmers to add the exact amount of nutrients needed in each field. The soil health card programme of

*Corresponding author: (Email: urwithnirmal@gmail.com)

Government of India has been embraced by all Indian states to provide soil health card to each farmer for sustainable crop production and as a first step toward precision farming. Thousands of data points on soil nutrients may now be used to create spatial maps of soil nutrients. Various researchers have utilized these data for soil nutrient mapping (Kusro *et al.* 2021; Singh *et al.* 2021). A map of the soil fertility in a particular location can be very helpful in determining how much fertilizer to use in that area (Kumar and Sinha 2018; Reza *et al.* 2021).

Several interpolation methods have been used to prepare soil nutrient maps from the point locations (Kumar 2013; Kumar 2018). Ordinary kriging has been found as one of the most extensively utilized interpolation methods for soil nutrient assessment. The procedure entails creating an empirical semivariogram, determining the best model fit for the semivariogram (*e.g.,* circular, spherical, Gaussian, exponential, and so on), and then kriging (Moharana *et al.* 2021; Sahu *et al.* 2020; Reza *et al.* 2021). Both grid based (Ayam *et al.* 2020) or random samples (Banwasi *et al.* 2020) have been used to prepare soil fertility maps. High density soil health card data have also been used to prepare soil fertility maps in various districts of Chhattisgarh including Kanker (Kusro *et al.* 2021), and Balod (Singh *et al.* 2021). This paper aims to prepare soil fertility maps of Bemetara district based on the high density soil health card data using ordinary kriging for site-specific nutrient management.

Materials and Methods

Study area

Bemetara district is 2854.81 km^{2}in size and is delimited by latitude 21°22' to 22°03' N and longitude 81°07' to 81°55 E (Fig. 1). The district is part of the Mahanadi Basin and has sedimentary deposit of the "Purana" basins (well-defined boundaries and not an erosional remnant or a tectonic depression) of Peninsular India (Karthikeyan *et al.* 2021). The district receives an average yearly rainfall of 1423.5 mm. A large acreage of cultivated land in the district is mainly under *kharif* crops such as rice (*Oriza sativa*), and soybean (*Glycine max*). The dominant *rabi* crop of the area is wheat (*Triticum aestivum*) grown with the support of irrigation and other crops like gram (*Cicer arietinum*) is cultivated on stored moisture. Few farmers grow lathyrus (*Lathyrus sativus*) and lentil (*Lens esculenta*) crops after the rice completely on stored moisture in the *rabi* season. Rice-wheat, rice-gram, soybean-gram and soybean-pigeon pea (intercrop) are most adopted cropping systems in the area.

Generation of spatial database of soil nutrient

Spreadsheet-formatted data from the soil health card was gathered and pre-processed to avoid duplications and positional errors. The spread sheet data were converted to spatial form by employing QGIS tools. A total of 14491 samples (Fig. 1) were divided in two parts, *i.e.* 70 % for training and 30 % for test.

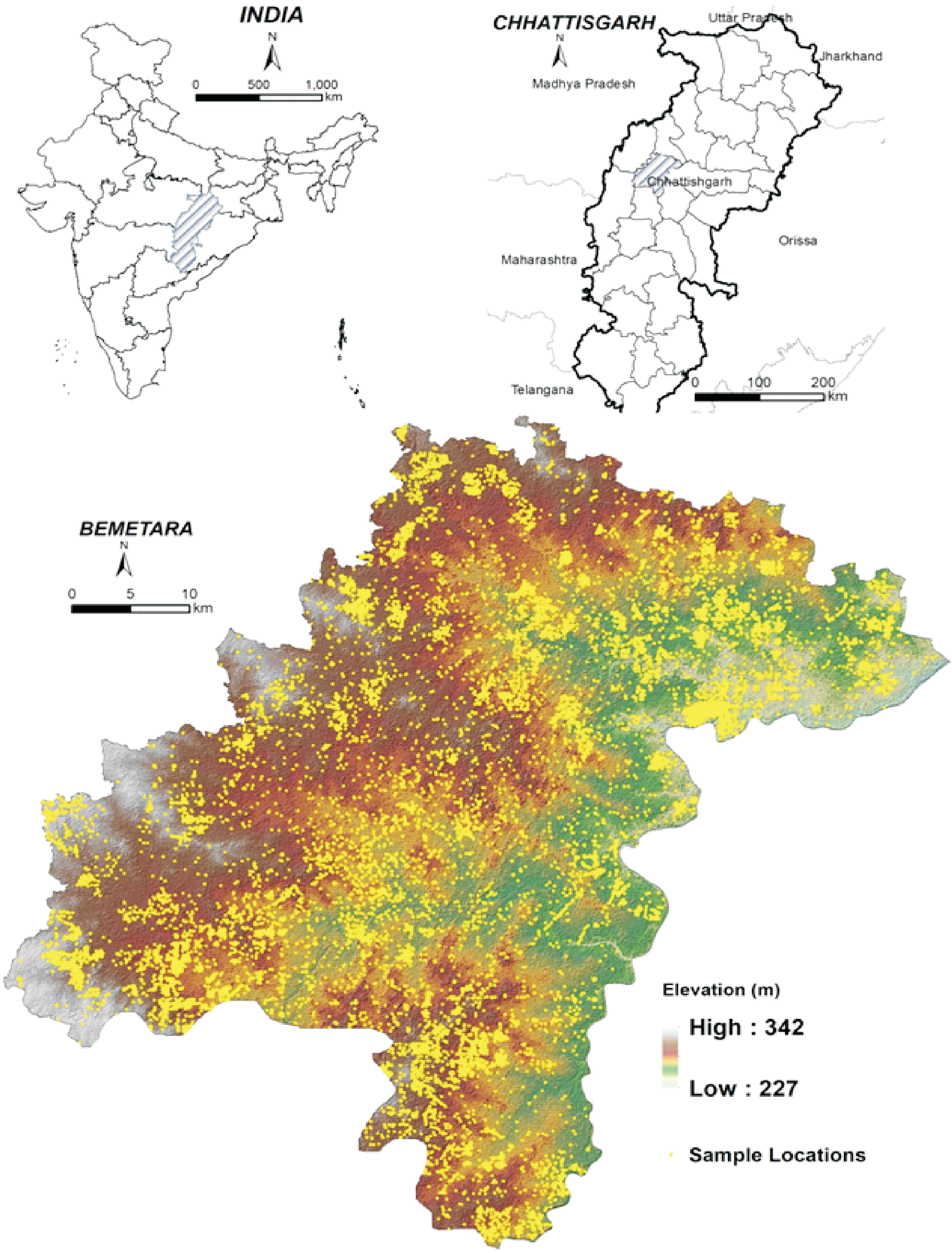


Fig. 1. Study area and sampling locations

Statistical and geo-statistical analysis

R (R core team 2020) - an open source software for statistics, predictive analytics, data visualization, and spatial analysis was used to process and analyze the soil nutrient data. This involves recognizing outliers, generation of training and validation sets, generation of descriptive statistics, and correlation analysis among the variables. Various packages in R were used as described by Kumar *et al.* (2021). The package '*metan*' (Olivoto and Lúcio 2020) was used for detection and removal of the outliers. The outliers were defined as values above third quartile plus inter-quartile range (IQR) or below first quartile minus IQR. The training and validation sets were divided in such a way that the descriptive statistics and the distribution of both the sets remain similar. The descriptive statistics including minimum, mean, median, maximum, standard deviation (SD), skewness, kurtosis and coefficient of variation (CV) were generated using the same package. Correlation matrix showing the correlations among the soil properties was generated using the package '*corrplot*' (Wei and Simko 2021).

Nutrient mapping

The ordinary kriging approach was used for mapping the soil properties (soil reaction, electrical conductivity, organic carbon, available nitrogen, phosphorous, and potassium). This includes generation of empirical semivariogram, model fitting, and interpolation. Semivariograms are used to represent spatial variability, implying homogeneity among equivalent lags to illustrate the mean deviation between observations differentiated by h. The values of the semivariogram at each lag separation (h) were measured:

$$\gamma(h) = 1/2(\square)\sum_{i=1}^{N(\square)}[z(xi) - z(xi + \square)]^2$$

Where,

$\gamma(h) = sample\ semi\,\mathrm{var}iance$

$N(h)$

= Numeric data combinations at a particular distance and direction class

$Z(x_i)$ =Value of variable at x_i point

$Z(x_i+h)$=Value of variable at a distance of h from the point x_i

Each soil attribute was fitted with one of four commonly used semivariogram models like circular, spherical, Gaussian, and exponential. The model with the lowest sum of squared error (SSE) was selected as the best fit model and its parameters (nugget, sill and range) were calculated (Webster and Oliver 2001). Nugget C_0- defines the micro-scale variability measurement error for the respective soil variable, sill (C) indicates the lag distance between measurements at which one value for a variable does not influence neighboring values and range (A) is the distance at which values of one variable become spatially independent of another (Lopez-Granados *et al.* 2002). ArcGIS Desktop ver. 10.2 was used for the geostatistical analyses.

Validation of kriged maps

The kriged maps were validated with the validation set to assess the authenticity of soil properties maps. The performance of each map was evaluated using R^2 (Pearson 1895), Lin's concordance correlation coefficient (LCCC) (Lin 1989), root mean square error (RMSE) (Kenney and Keeping 1962) and bias (Varma and Simon 2006).

Results and Discussion

Descriptive statistics

The descriptive statistics are shown in the table 1. electrical conductivity (EC), organic carbon (OC), available nitrogen (N), phosphorous (P), and potassium (K) were 0.37 dSm^{-1}, 0.69 %, 228.5 kg ha^{-1}, 17.24 kg ha^{-1} and 390 kg ha^{-1}, respectively. The skewness and kurtosis values were well within the range of -1 to +1, indicating normal distribution of the data and no substantial skewness and peakness in the data. Coefficient of variation (CV) is one of the most important parameter that describes the variability of soil variables than other parameter such as standard deviation (SD), mean and median (Xing-Yi *et al.* 2007 and Zhou *et al.* 2010). Hence, the variability of soil variables was interpreted as per Wilding (1985) using the CV classes as highly variable (CV > 35%), moderately variable (CV 15 to 35%) and low variable (CV < 15%). Accordingly, high variability was found for EC while the other parameters were moderately variable.

Table 1. Statistical summary of soil parameters

Variable	Min	Max	Mean	Median	Standard Deviation	Kurtosis	Skewness	Coefficient of variation
EC	0.02	0.84	0.37	0.33	0.15	0.50	1.09	42.459
OC	0.11	1.29	0.69	0.7	0.21	-0.12	-0.02	31.614
N	100.35	365.6	228.50	225.79	49.36	-0.45	0.00	21.605
P	0.89	33.92	17.24	17.02	5.78	-0.24	0.24	33.562
K	88.4	692.99	390.09	378.5	110.14	-0.32	0.41	28.235

Correlation among soil nutrients

The relationships among the soil parameters in the form of a correlation plot is shown in figure 2. The negative and positive significant correlations are coloured red and blue, respectively. The strength of the connections is indicated by the size and intensity of the colour. The correlations were moderate. The corrplot showed insignificant correlations between pH and EC and between N and P at significance level of 95%. Theother correlations were significant. The weak correlation was shown in fade colours and small circles.

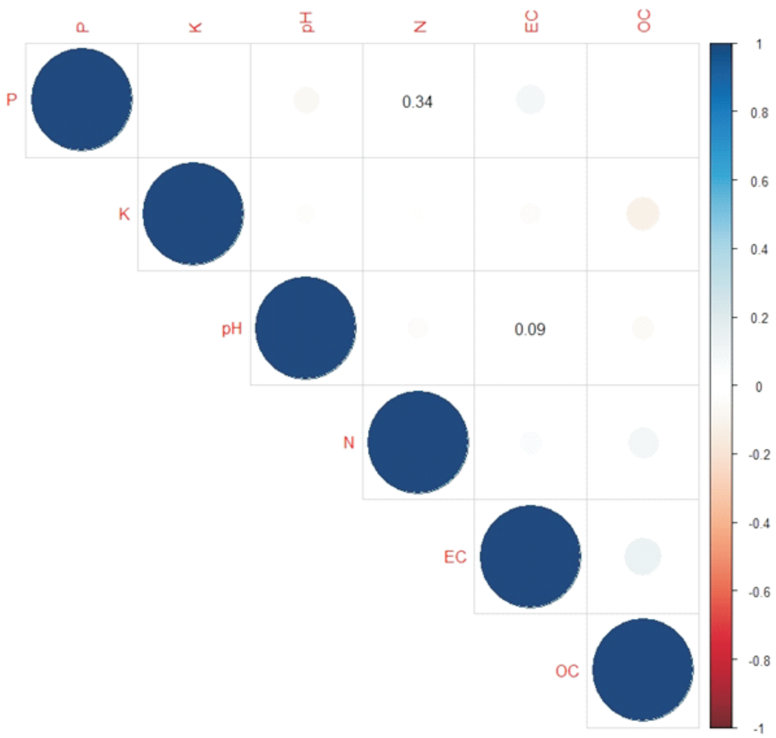


Fig. 2. Correlation matrix for the soil nutrient parameters

Training and test data

The total sample was divided into two parts based on the variability of the soil properties. 10144 samples were identified as training and 4347 samples as test. The distribution of the data for all parameters is shown in figure 3. The descriptive statistics of both the data sets were found similar in all variables. The figure showed that the mean, median, the quartiles and the inter-quartile range (IQR) were same for the training and test datasets for all the soil properties.

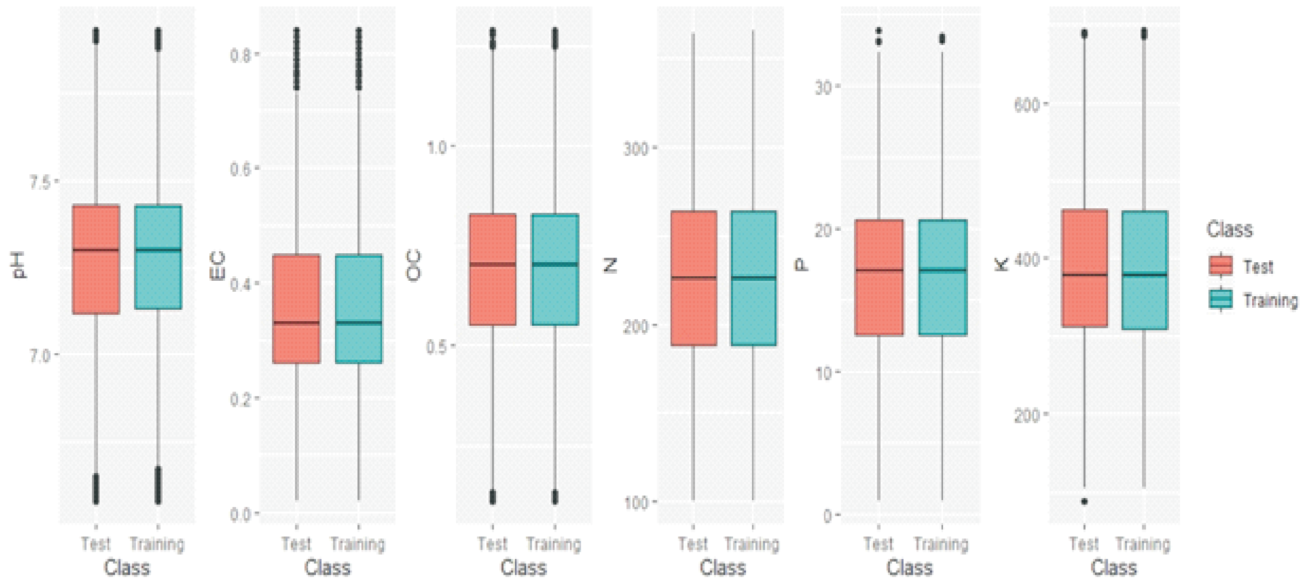


Fig. 3. Distribution of training and test data

Modeling spatial variability

To determine the various spatial patterns of distinct soil nutrients, semivariograms were created for each parameter. (Fig. 4) The best fit models for all variables using semivariogram parameters (range, nugget, and partial sill) is shown in table 2. From the plotted graphs of semivariogram and after computing nugget/sill ratio, there is weak spatial dependency incase of EC, pH, and K. Moderate dependency found in case of OC and N, whereas, strong spatial dependency was found in case of P. Nugget/sill of higher ratio indicated that the spatial variability was primarily caused by stochastic factors *e.g.* fertilization, farming measures, cropping systems and other human activities. A lower nugget /sill ratio showed that the structural factors *e.g.* factors of soil genesis play significant role in spatial variability. Thus, the OK method of interpolation do not provide better prediction for soil nutrients at unknown points as observed in figure 5.

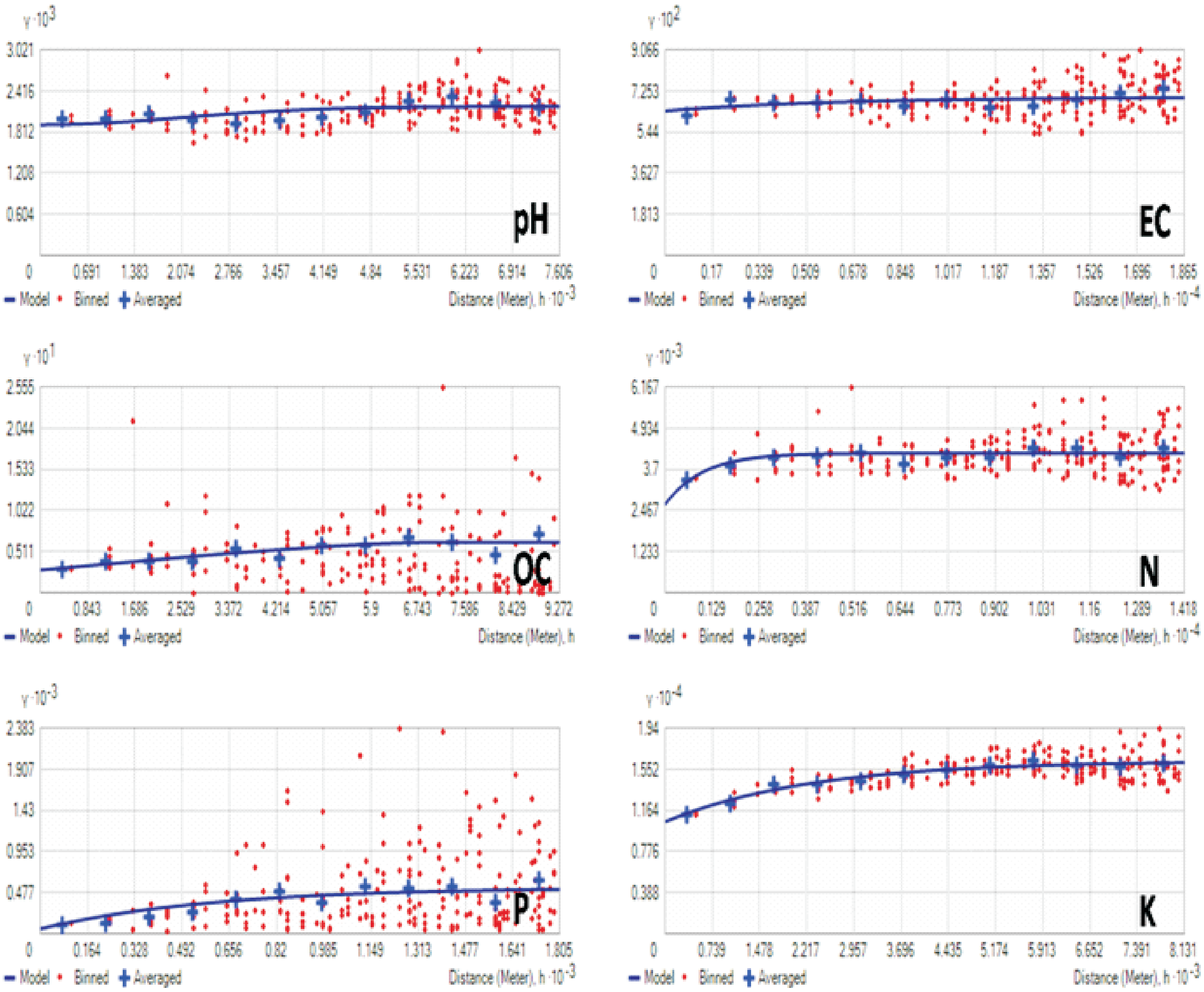


Fig. 4. Semivariogram model for the soil parameters

Table 2. Best fit model for the empirical semivariograms and their statistics

Parameters	**Model**	**Nugget**	**Partial sill**	**Sill**	**Lag**	**Range**	**Nugget/ sill**	**Spatial Dependency**
pH	Gaussian	0.002	0.0003	0.0023	633.8	7605	0.86	Weak
EC	Exponential	0.063	0.006	0.0690	1554	18652	0.91	Weak
Organic carbon	Circular	0.028	0.0342	0.0622	0.77	6.85	0.44	Moderate
Avail Nitrogen	Exponential	2647	1534.7	4181.7	1181	3084	0.63	Moderate
Phosphorous	Exponential	59.5	481.55	541.05	150	1805	0.10	Strong
Potassium	Exponential	10970	4035	15005	677	6862	0.73	Weak

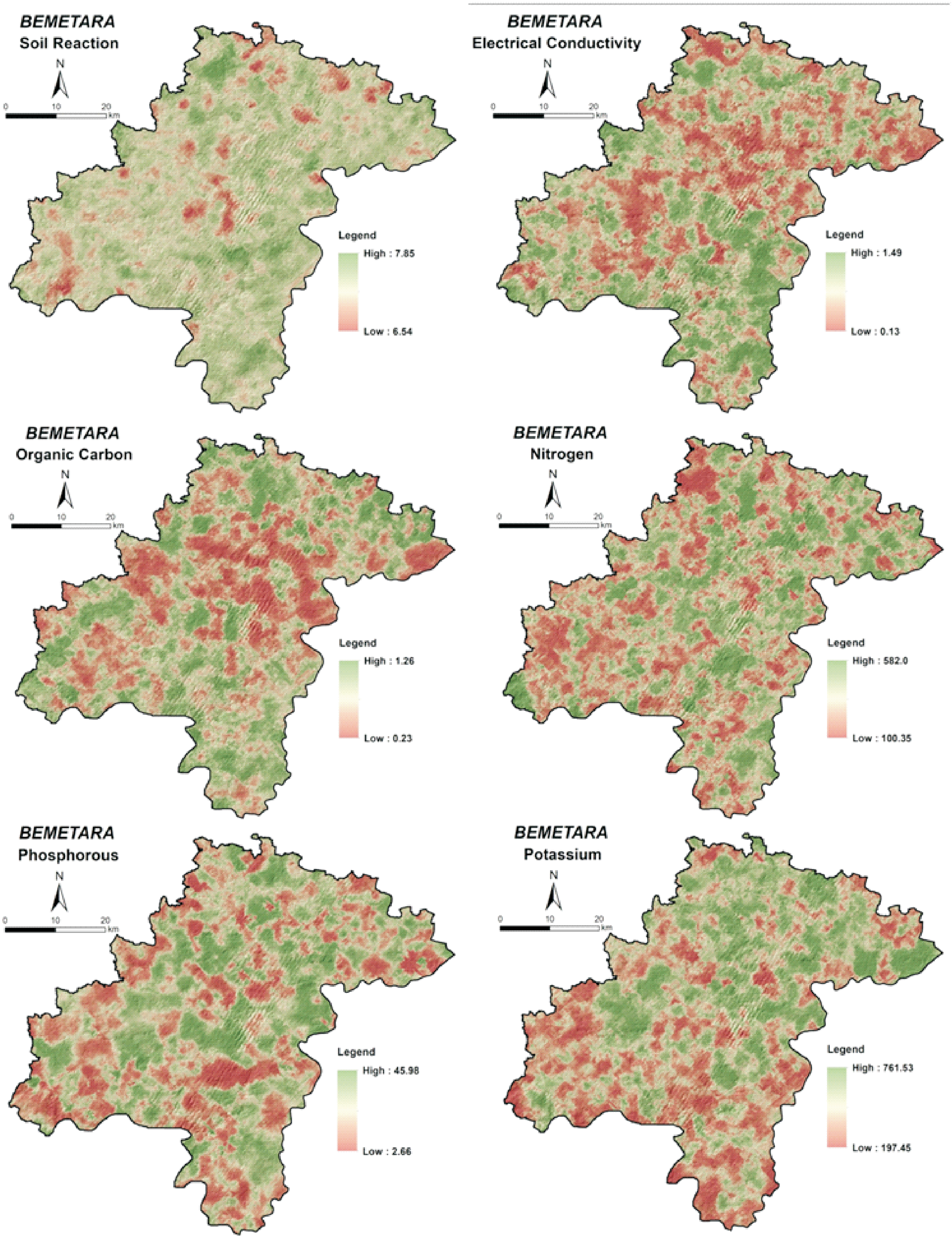


Fig. 5. Ordinary kriging maps for the soil nutrient variables (pH, EC, OC, N, P, and K)

The soil reaction ranged from 6.5 to 7.8 with lower pH in the forest and hilly areas. Similarly, the soil EC ranged from 0.13 to 1.49 dS m^{-1} with lower values in the forest areas and higher in case of the plains with agriculture. The relationships between soil pH and terrain indicators such as slope are well established (Moore *et al.* 1993; Chen *et al.* 1997; Li *et al.* 2017; Ayam *et al.* 2020). The lower pH values in forests can be attributed to accumulation and subsequent slow decomposition of organic matter, which releases acids. Organic carbon found in the range of 0.23-1.26 % while higher carbon percentage is found in the forest region. Nitrogen ranged from low (100 kg ha^{-1}) to high (585 kg ha^{-1}) with lower values in the hilly areas. Phosphorous and Potassium ranged from 2.66 to 46 kg ha^{-1} and 197.45 to 761.6 kg ha^{-1} respectively.

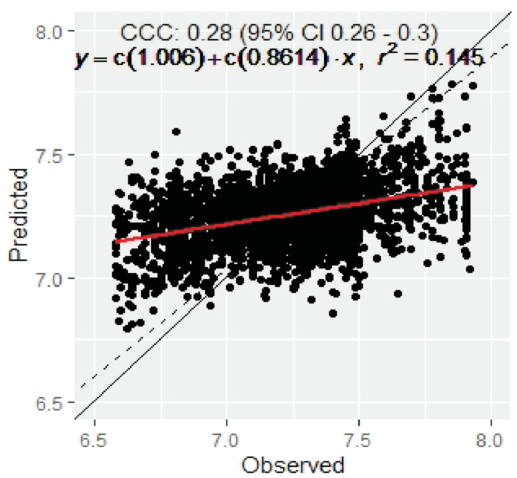
Validation and accuracy assessment

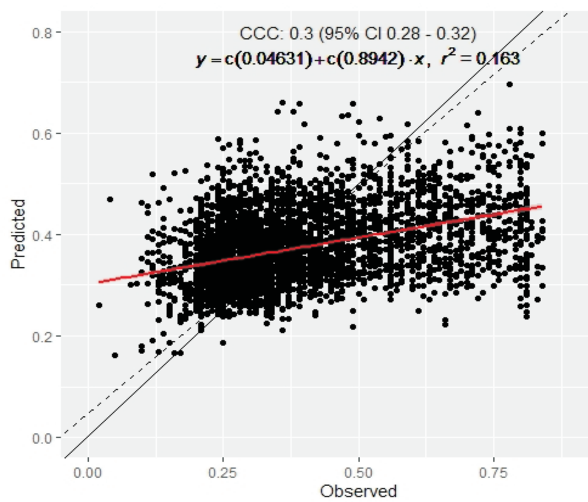
Spatial distribution maps of soil properties were validated using the validation set of 4347 sample points. Evaluation indices of the predicted pH, EC, OC, N, P, and K through OK approach are shown in table 3. The observed versus average predicted values are shown in fig. 6.

Table 3. Validation indices in accuracy of spatial analysis

Parameter	RMSE	R^2	P Value	LCCC	Bias
pH	0.231	0.145	0.000	0.28	-0.0005
EC	0.147	0.163	0.000	0.30	0.0070
OC	0.201	0.158	0.000	0.29	0.0003
N	47.980	0.050	0.000	0.14	0.1154
P	5.479	0.096	0.000	0.20	-0.0402
K	98.15	0.208	0.000	0.37	1.6991

Evaluation indices for the predicted maps showed the RMSE values of N and P were high as compared to other parameters (Table 3). The R^2 values for all soil variables were poor but significant. The LCCC values ranged from 0.14 to 0.37 for different soil properties in the area showing a poor correlation. Overall, the bias of the prediction was positive for EC, OC, N, and K, and negative for pH and P.





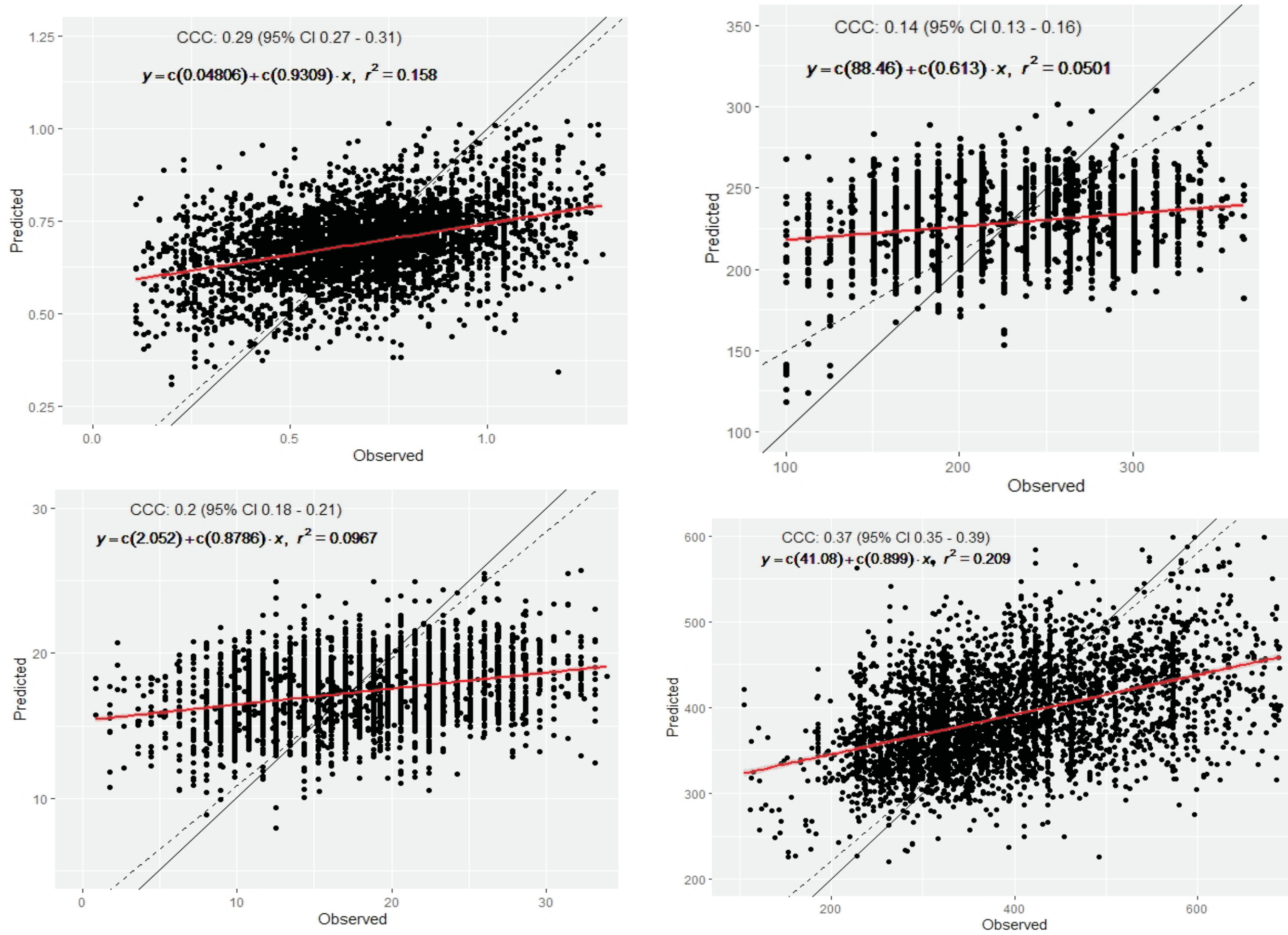


Fig. 6. Validation statistics between observed and predicted soil parameters

Conclusion

Soil health card data can be used for generating high resolution maps utilizing spatial interpolation techniques that can convert geographically discrete data to continuous surface maps and ultimately reduce the number of sampling sites as a result cumbersome sampling procedure can be averted. Ordinary kriging is most widely used form of kriging but the maps generated though it is less accurate. Inclusion of the factors of climate, land use, and terrain which influence soil formation and properties of soil effectively for prediction of soil properties through machine learning techniques such as random forest, support vector machine, *etc.* may result in better modeling accuracies.

References

Ayam, G. P., Obi Reddy G.P., Kumar, N., Singh, S.K., Sahu, K.K., Shrivastava, G.K., Saxena, R.R. and Deshmukh, B.L. (2020). Characterization of landform in Bastar plateau using high resolution ALOS-DEM data. *International Journal of Chemical Studies* **8,** 997-1000.

Banwasi, R., Singh, S.K., Kumar, N. and Sahu, K.K. (2020). Assessing the accuracy of geostatistical techniques for mapping soil macronutrients on basaltic landscape of central India. *International Journal of Chemical Studies* **8,** 1014-1019.

Chen, Z. S., Hsieh, C. F., Jiang, F. Y., Hsieh, T. H. and Sun, I. F. (1997). Relations of soil properties to topography and vegetation in a subtropical rain forest in southern Taiwan. *Plant Ecology* **132**, 229–241.

Karthikeyan, K., Padikkal, C., Tiwary, P. Vasu, D., Paul, R., Kumar, N., Anatwar, S., Kuchankar, H., Gautam, N., Sahu, V. and Gaikwad, S. (2021). Shrink–swell soils on shale geology of Chhattisgarh Basin, India: its formation and significance. *Arabian Journal of Geosciences* **14**, 1059.

Kenney, J. F. and Keeping, E.S. (1962). Root Mean Square. In 'Mathematics of Statistics', Part 1, (3rd Edition). Van Nostrand, Princeton, pp. 59–60.

Kumar, N. (2018). Study on identification, characterization and mapping of degraded lands using time series MODIS NDVI and LANDSAT data. A dissertation for doctoral degree award in Indira Gandhi Krishi Vishvavidyala, Raipur, 256.

Kumar, N., Sinha, N.K., Sahoo R.N. and Santra, P. (2021). Manipulating and preprocessing data with R. Statistics with R series. Brillion Publishing. 253.

Kumar, N. and Sinha, N.K. (2018). Geostatistics: Principles and Applications in Spatial Mapping of Soil Properties. In 'Geospatial Technologies in Land Resources Mapping, Monitoring and Management'. (Eds. Reddy G., Singh S.). Geotechnologies and the Environment, Volume 21. Springer, Cham.

Kumar, N. (2013). Geostatistics in digital terrain analysis. Remote sensing and GIS in digital terrain analysis and soil-landscape modeling, 152.

Kusro, P.S., Tedia, K., Bajpai, R.K., Sonboir, H.L. and Saxena, R.R. (2021). Geoinformatics for evaluating and mapping soil fertility index in a rice growing area of Kanker district in Chhattisgarh. A dissertation for doctoral degree award in Indira Gandhi Krishi Vishvavidyala, Raipur, 230.

Lin, L. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics* **45**, 255–268.

Li, X., Chang, S. X., Liu, J., Zheng, Z., Wang, X. (2017). Topography-soil relationships in a hilly evergreen broadleaf forest in subtropical China. *Journal of Soils and Sediments* **17,** 1101–1115.

Lopez-Granados, F., Jurado-Exposito, M., Atenciano, S., Gracia-Ferrer, A., De La Orden, M. S., Gracia-Toreres, L. (2002). Spatial variability of agricultural soil parameters in Southern Spain. *Plant Soil* **246,** 97–105.

Moharana, P. C., Jena, R. K., Kumar, N., Singh, R. S. and Rao, S.S. (2021). Assessment of soil organic and inorganic carbon stock at different soil depths after conversion of desert into arable land in the hot arid regions of India. *Carbon Management* **12,** 153-165.

Moore, I.D., Gessler, P.E., Nielsen, G.A.E. and Peterson, G.A. (1993). Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal* **57**, 443–452.

Olivoto, T. and Lúcio, A.D. (2020). metan: an R package for multi-environment trial analysis. *Methods in Ecology and Evolution* **11,** 783-789

Pearson, K. (1895). Notes on regression and inheritance in the case of two parents. *Proc Royal Soc London* **58,** 240–242.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reza, S.K., Kumar, N., Ramachandran, S., Mukhopadhyay S., Singh S. K., Dwivedi B. S. and Ray, S.K. (2021). Geo-spatial analysis for horizontal and vertical variability of bulk density, particle-size distribution and soil moisture content in Tripura, Northeastern India. *Arabian Journal of Geosciences* **14**, 2734.

Sahu, N., Obi Reddy G.P., Kumar, N., Nagaraju, M.S.S., Srivastava, R. and Singh, S.K. (2020). Spatial variability of soil macronutrients on basaltic landscape of Central India: A geostatistical approach. *Chemical Science Review and Letters* **9,** 77-86.

Singh, A., Mishra, V. N., Kumar, N., Tedia, K., Pandey, N. and Saxena, R. R. (2021). Nutrient mapping of Balod district of Chhattisgarh with machine learning technology using soil health card data. A dissertation for post graduation degree award in Indira Gandhi Krishi Vishvavidyala, Raipur, 101.

Varma, S. and Simon, R. (2006). Bias in error estimation when using cross validation for model selection. BMC Bioinformatics **7**, 91.

Webster, R. and Oliver, M. A. (2001). Local estimation or prediction: kriging. Geostatistics for environmental scientists. Wiley, Chichester, pp. 149–191.

Wei, T. and Simko, V. (2021). R package 'corrplot': Visualization of a Correlation Matrix (Version 0.92).

Wilding, L.P. (1985). Spatial Variability: Its Documentation, Accommodation, and Implication to Soil Surveys. In 'Soil spatial variability'. (Eds. Nielsen, D.R. and Bouma, J.) Pudoc, Wageningen, Netherlands, pp. 166- 194.

Xing-Yi, S.Y.Y., Zhang, X.D., Meng, K. and Herbert, S.J. (2007). Spatial variability of nutrient properties in black soil of Northeastz China. *Pedosphere* **17**, 19–29.

Zhou, H., Chen, Y. and Li, W. (2010). Soil properties and their spatial pattern in an oasis on the lower reaches of the Tarim River, Northwest China. *Agricultural Water Management* **97**, 1915–1922.

Received: September, 2021 *Accepted: April, 2022*