Journal of the Indian Society of Soil Science, Vol. 60, No. 2, pp 108-115 (2012)

Spatial Variability of Soil Properties in Brahmaputra Plains of North-eastern India: A Geostatistical Approach

S.K. Reza*, Utpal Baruah and Dipak Sarkar[1]

National Bureau of Soil Survey and Land Use Planning, Regional Centre, Jorhat, 785 004, Assam

The objective of this study was to determine, the degree of spatial variability of pH, organic carbon (OC), available nitrogen (AN) and available potassium (AK) of Goalpara district of Assam. A total of 1397 soil samples from a depth of 0-25 cm at an approximate interval of 1 km were collected over the entire district. Data were analyzed both statistically and geostatistically on the basis of semivariogram. Soil properties showed large variability with greatest variation observed in AK (42%) whereas the smallest variation was in pH (16%). The semivariogram for all soil properties were best fitted by exponential models and showed a highest (3.8 km) range for pH and lowest (2.0 km) for AN. The nugget/sill ratio indicates a strong dependence for pH (19%), moderate spatial dependence for OC (60%) and a weak spatial dependence for other soil properties. Evaluation of spatial maps indicated that except for AN, kriging could successfully interpolate other soil properties. Soil pH was significantly negatively correlated with OC (r = -0.434**) and AN (r = -0.228**). A significant positive relationship was observed between OC with AN (r = 0.498**) and AK (r = 0.119**).

Key words: Spatial variability, kriging, semivariogram, accuracy assessment, soil properties, Brahmaputra plains

Site-specific management has received considerable attention due to potential benefits of increasing input use efficiency, improving the economic margins of crop production and reducing environmental risks (Yasrebi *et al.* 2008). Hence, a comprehensive understanding of spatial variability of soil properties is becoming increasingly essential in agriculture. Soil properties vary spatially from field to a large region scale and are influenced by geology, topography, climate as well as soil use (Quine and Zhang 2002). In addition, variability can occur as a result of land use and management strategies (Wang *et al.* 2009).

Developing accurate application maps for site-specific fertilization is critical in implementing precision farming technology. Therefore, spatial variability map showing soil properties (pH, organic carbon, nitrogen and potassium) will make it possible to reduce fertilizer use, costs and environmental pressure (Lopez-Granados *et al.* 2002). Soil surveys provide a map of different soil orders, together with a record of measured observations for each sampling location. However, for a large-scale soil property map it is not practicable to measure soil properties for all the locations. Kriging is a useful tool to predict and interpolate data between measured locations (Burgess and Webster 1980).

In recent years, geostatistics has proved to effective in assessing the variability of soil nutrients (Webster and Oliver 2001; Gilbert and Wayne 2008). In India geostatistics has been widely used by many workers for spatial variability of soil properties like for phosphorus (Grewal *et al.* 2001), salinity (Nayak *et al.* 2002), boron (Chinchmalatpure *et al.* 2005), micronutrients (Nayak *et al.* 2006) and soil properties and hydraulic parameters (Santra *et al.* 2008). Geostatistics is a technology for estimating the soil property values in nonsampled areas or areas with sparse samplings (Yao *et al.* 2004). These nonsampled areas can vary in space (in one, two or three dimensions) from the sampled data (Zhu *et al.* 2005). Geostatistics provides a set of statistical tools for a description of spatial patterns, quantitative modeling of spatial continuity, spatial prediction, and uncertainty assessment (Goovaerts 1999). Geostatistical techniques incorporating spatial information into predictions can improve estimation and enhance map quality (Mueller and Pierce 2003).

*Corresponding author (Email: reza_ssac@yahoo.co.in)
Present address
[1] National Bureau of Soil Survey and Land Use Planning, Nagpur, 440 030, Maharashtra

Most soil variability studies have been carried out in hot-arid or semi-arid climates in the country (Nayak *et al.* 2006; Santra *et al.* 2008). Although these studies provided very precise information for site-specific recommendations, for more precision, it is necessary to consider the fact that spatial variability of soils depends on the specific soil studied. In other words, the organic matter in the same field may be fairly static; soil mineral nitrogen may be highly variable over time as well as in space (Geypens *et al.* 1999). A few reports have been obtained to understand the spatial variability of the soil properties in humid subtropical India (Reza *et al.* 2010). Thus, the objective of this study was to determine the spatial variability of selected soil properties such as pH, organic carbon content (OC), available N (AN), and available K (AK), with the classical statistics and geostatistical analysis for Goalpara district of Assam, under Brahmaputra plains.

Materials and Methods

Site Characteristics

The area under study belongs to the Goalpara district of Assam (25º53"-26º30"N, 90º07"-91º05"E) covering an area 1911 km^2 (Fig. 1) of Brahmaputra plains. The topography of the district is generally characterized by an almost flat plain except for few low-forested hills that break the monotony of the terrain. The climate in the district is moderate during the winter and in summer it is hot. The maximum temperature is 33 ºC during July and August; a minimum temperature falls up to 7 ºC in the month of January. Rain makes its first appearance in the month of April with occasional and irregular light showers and at times, heavy down pour followed by cyclonic storm. Annual rainfall is 2169 mm and about 80% of rainfall is from the south-west monsoon. Significance of the district is the existence of a large number of

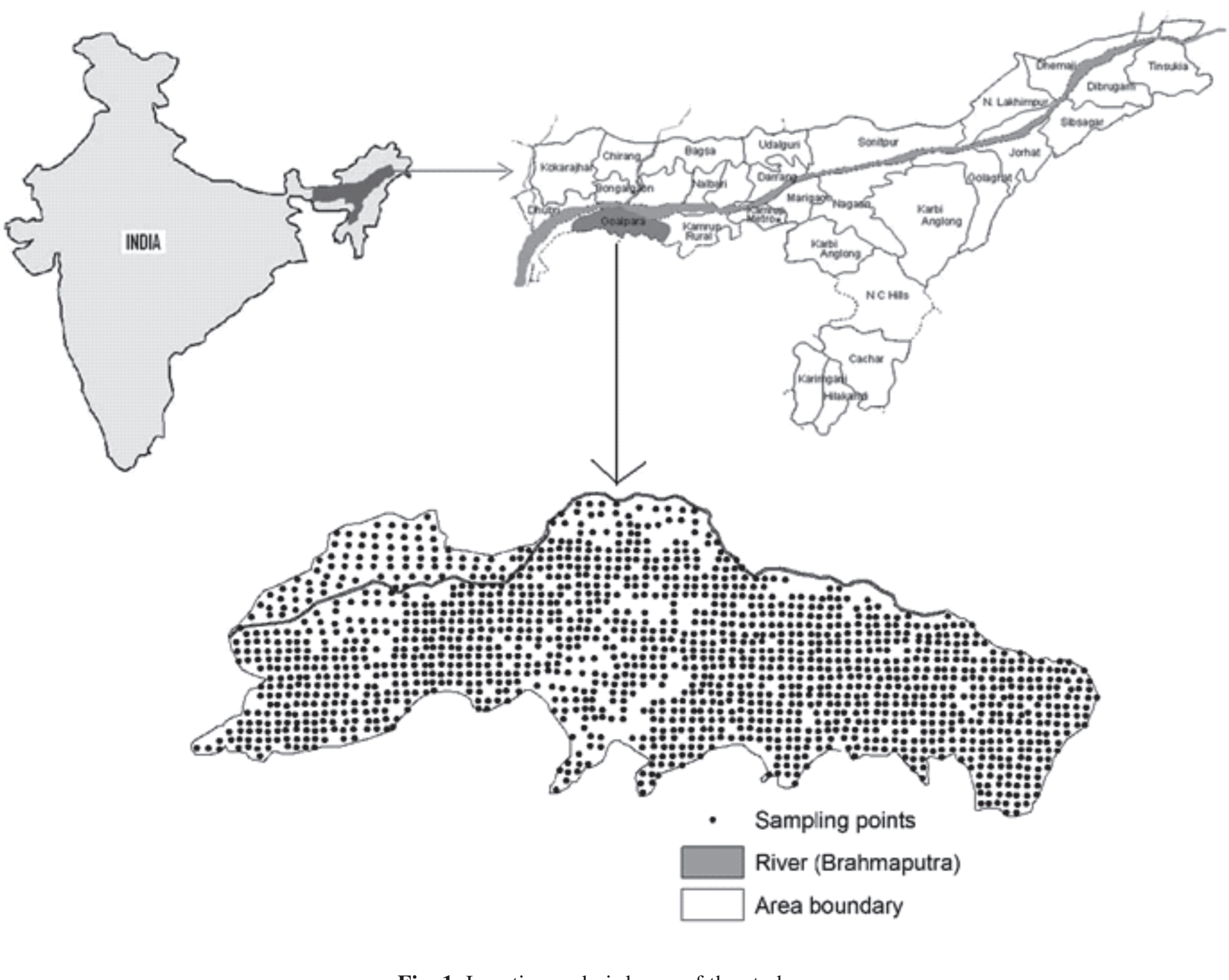

Fig. 1. Location and gird map of the study area

char (riverine tracts and sandy river island) in the river Brahmaputra. There are eight broad soil sub-groups in the district according to Soil Taxonomy (USDA) namely-Aeric Fluvaquents, Aeric Haplaquepts, Aeric Haplaquents, Typic Udifluvents, Typic Kandihumults, Typic Haplumbrepts, Dystric Eutrochrepts and Typic Paleudults (NBSS&LUP 1999).

Soil Sampling, Processing and Analysis

In autumn of 2008, a total of 1397 soil samples were collected from the plough layer (0-25 cm) at an approximate interval of 1 km grid (Fig. 1) with the help of hand-held global positioning system (GPS) over the entire Goalpara district of Assam. Soil samples were air-dried and ground to pass through a 2-mm sieve. Available K was extracted with 1*M* NH_4OAc and then measured by atomic absorption spectrophotometer. Available N was determined by Subbiah and Asija (1956), organic carbon by Walkley and Black (1934) and pH with glass electrode in a 1:2.5 soil/water suspension.

Statistical and Geostatistical Analysis

The main statistical parameters, including mean, standard deviation, variance, coefficient of variance, and extreme maximum and minimum values, which are generally accepted as indicators of the central tendency and of the data spread, were analyzed. The Pearson correlation coefficients were estimated for all possible paired combinations of the response variables to generate a correlation coefficient matrix. These statistical parameters were calculated with EXCEL® 2007 and SPSS 15.0® (SPSS Inc., Chicago, III., USA).

Geostatistics (Matheron 1963) uses the semivariogram to quantify the spatial variation of a regionalized variable, and provides the input parameters for the spatial interpolation method of Kriging (Krige 1951). The semivariogram is half the expected squared difference between paired data values $z(x)$ and $z(x+h)$ to the lag distance, by which locations are separated (Webster and Oliver 2001):

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2 \qquad \ldots(1)$$

where, $z(x_i)$ is the value of the variable z at location of x_i, h the lag and $N(h)$ the number of pairs of sample points separated by h. For irregular sampling, it is rare for the distance between the sample pairs to be exactly equal to h. That is, h is often represented by a distance band.

Spherical, exponential and Gaussian models were fitted to the empirical semivariograms. Best-fit model with smallest nugget values with minimum root mean square error (RMSE) were selected for each soil property. The expression of RMSE is given below.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[z(x_i) - \hat{z}(x_i)]^2} \qquad \ldots(2)$$

Expression for exponential semivariogram model best fitted in this study is given below.

$$\gamma(h) = C_o + C_1\left[1 - exp\left\{-\frac{h}{a}\right\}^2\right] \; for\; h \geq 0 \qquad \ldots(3)$$

Using the model semivariogram, basic spatial parameters such as nugget, sill and range was calculated which provide information about the structure as well as the input parameters for the kriging interpolation. Nugget is the variance at zero distance, sill is the lag distance between measurements at which one value for a variable does not influence neighboring values and range is the distance at which values of one variable become spatially independent of another (Lopez-Granadoz *et al*. 2002). Geostatistical analysis consisting of semivariogram calculation, cross-validation and mapping was performed using the geostatistical analyst extension of ArcGIS 9.3.1 version (ESRI 2009).

Accuracy of the soil maps was evaluated through cross-validation approach (Davis 1987). Among three evaluation indices used in this study, mean absolute error (MAE), and mean squared error (MSE) measure the accuracy of prediction, whereas goodness of prediction (*G*) measures the effectiveness of prediction. The MAE is a measure of the sum of the residuals (*e.g.* predicted minus observed) (Voltz and Webster 1990).

$$MAE = \frac{1}{N}\sum_{i=1}^{N}[\![z(x_i) - \hat{z}(x_i)]\!] \qquad \ldots(4)$$

where, $\hat{z}$ is the predicted value at location *i*. Small MAE values indicate less error. The MAE measure, however, does not reveal the magnitude of error that might occur at any point and hence MSE will be calculated,

$$MSE = \frac{1}{N}\sum_{i=1}^{N}[z(x_i) - \hat{z}(x_i)]^2 \qquad \ldots(5)$$

Squaring the difference at any point gives an indication of the magnitude, *e.g.* small MSE values indicate more accurate estimation, point-by-point. The *G* measure gives an indication of how effective a prediction might be relative to that which could have been derived from using the sample mean alone (Schloeder *et al.* 2001).

$$G = \left[1 - \frac{\sum_{i=1}^{N}[z(x_i) - \hat{z}(x_i)]^2}{\sum_{i=1}^{N}[z(x_i) - \bar{z}]^2}\right] \times 100 \quad \ldots(6)$$

where, z is the sample mean. If $G = 100$, it indicates perfect prediction, while negative values indicate that the predictions are less reliable than using sample mean as the predictors. The comparison of performance between interpolations was achieved by using mean absolute error (MAE).

Results and Discussion

Descriptive Statistic

The summary of the statistics for soil properties are shown in table 1. The median of each soil properties was lower than the mean, which indicates that the effects of abnormal data on sampling value were not great. There was difference in the CV of the soil properties. The greatest variation was observed in AK (42%) where as the smallest variation was in pH (16%). Other researchers also documented a smaller variation of soil pH compared to other soil properties (Tsegaye and Hill 1998; Sun *et al.* 2003). This may be attributed to the fact that pH values are log scale of proton concentration in soil solution, there would be much greater variability if soil acidity is expressed in terms of proton concentration directly. All the soil properties exhibit a medium variation (15-50%) according to guidelines provided by Warrick (1998). Skewness is the most common form of departure from normality. If a variable has positive skewness, the confidence limits on the variogram are wider than they would otherwise be and consequently, the variances are less reliable. A logarithmic transformation is considered where the coefficient of skewness is greater than one (Webster and Oliver 2001). Therefore, a logarithmic transformation was performed for pH and AK parameters as their skewness was greater than 1.

Geostatistical Analysis

Spatial dependence of soil properties: In order to identify the possible spatial structure of different soil properties, semivariograms were calculated and the best model that describes the spatial structure was identified based on minimum RMSE (Table 2). The optimal theoretical model for all the soil properties was best fitted to the exponential models. The range expressed as distance could be interpreted as the diameter of the zone of influence that represented the average maximum distance over which a soil property of two samples was related. At distances less than the range, measured properties of two samples became similar with decreasing distance between the two points. Thus, the range provided an estimate of areas of spatial-dependence. The zones of influence for pH and OC were 3.8 and 2.9 km, respectively; however, for AK and AN they were smaller, being only 2.3 and 2.0 km, respectively (Table 2); thus, the length of the spatial autocorrelation is much longer than the sampling interval of 1 km. Therefore, the current sampling design is appropriate for this study and it is expected that a good spatial structure will be shown on the interpolated map. All soil properties showed positive nugget, which can be explained by sampling error, short range variability, random and inherent variability. To define different classes of spatial dependence for the soil variables, the ratio, the nugget and sill was used (Cambardella *et al.* 1994). The variable is considered to have a strong spatial dependence if the ratio is less than 25%, and has a moderate spatial dependence if the ratio is between 25 and 75%; otherwise, the variable has a weak spatial dependence. The nugget/sill ratio showed a strong spatial dependence for pH (19%), which might be attributed to the strong leaching process of soil cations in this subtropical region. A similar result was also reported by Sun *et al.* (2003) in the hill region of subtropical China. Organic carbon (60%) showed a moderate spatial dependence, which might be attributed to extrinsic factor (fertilization and cultivation practices) and intrinsic (soil-forming processes), whereas AN and AK were weakly spatially dependent (Table 2).

Spatial distribution of soil properties: The parameters of the exponential model were used for

Table 1. Descriptive statistics for pH, organic carbon (OC), available nitrogen (AN) and available potassium (AK) in soils of Goalpara district, Assam (n=1397)

Parameters	Min.	Max.	Mean	Median	SD*	CV (%)**	Skewness	Kurtosis
pH	3.5	7.8	4.9	4.8	0.81	16	1.33	4.69
OC (g kg^{-1})	0.4	36.3	13.7	13.4	5.28	38	0.23	3.30
AN (mg kg^{-1})	22.7	387.0	191.4	183.0	65.93	34	0.11	2.55
AK (mg kg^{-1})	7.0	191.5	52.4	47.3	21.86	42	1.40	5.84

*Standard deviation; **Coefficient of variation

Table 2. Models and parameters of semivariograms of soil properties

Soil properties	Fitted model	Nugget	Sill	Range	Nugget/Sill (%)	RMSE*
pH	Exponential	0.007	0.036	3.8	19	0.487
OC	Exponential	18.18	30.10	2.9	60	4.468
AN	Exponential	3613.3	4352.5	2.0	83	64.26
AK	Exponential	0.132	0.166	2.3	79	21.10

*Root mean square error

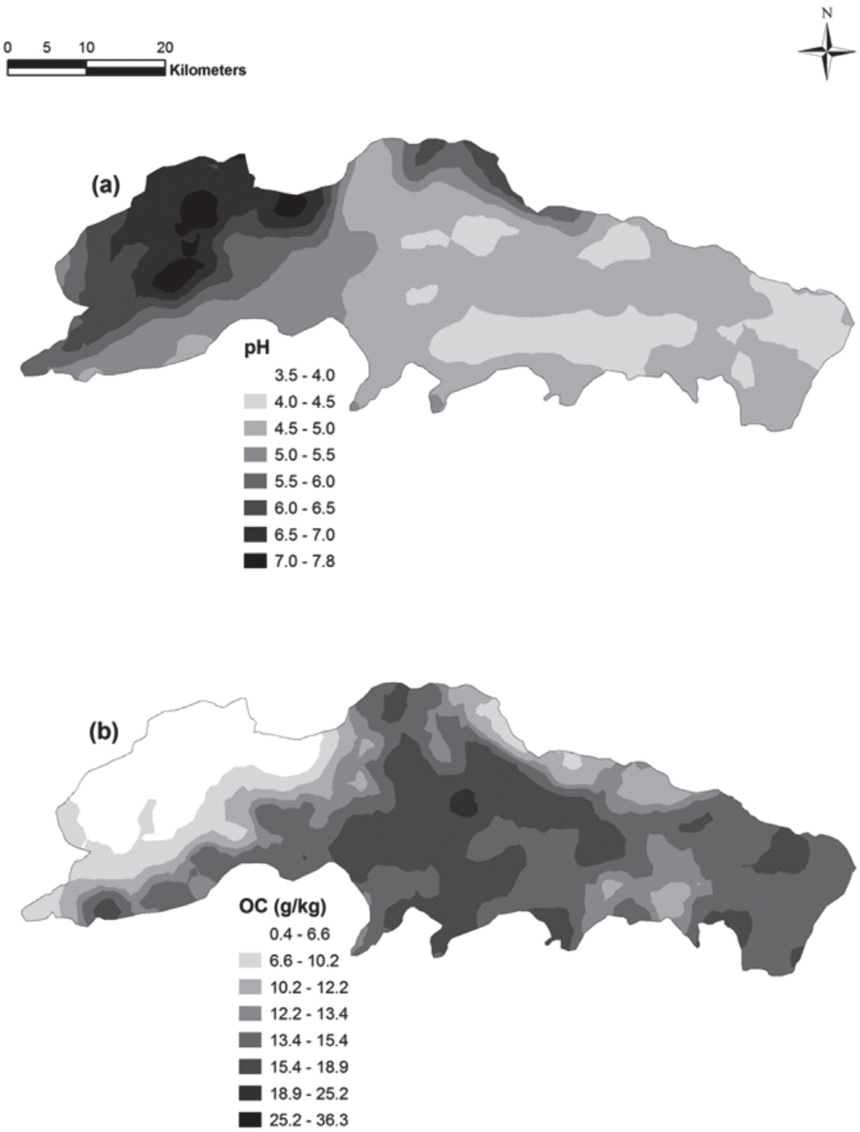

Fig. 2. Maps of the kriged estimates for (a) pH and (b) organic carbon (OC) (g kg^{-1})

kriging to produce the spatial distribution maps of soil properties of the study area. A search region of 12 nearest neighbours was applied. Spatial maps of pH and OC (Fig. 2) and, AN and AK (Fig. 3) prepared through kriging showed that pH value in the western part of the study area was in neutral to alkaline range along with the Brahmputra river and the value in the southeast quadrant was strongly acidic (< 4.5 pH) as this part of the district has coarse to medium grained, massive, grey to pink granites which is the extension of Meghalaya plateau. Organic carbon and AN had a large similar spatial variability; both these soil properties decreased in the western part of the study area and increased in central, southwest and southeast quadrant, while AK had an inverse distribution may be due to landscape.

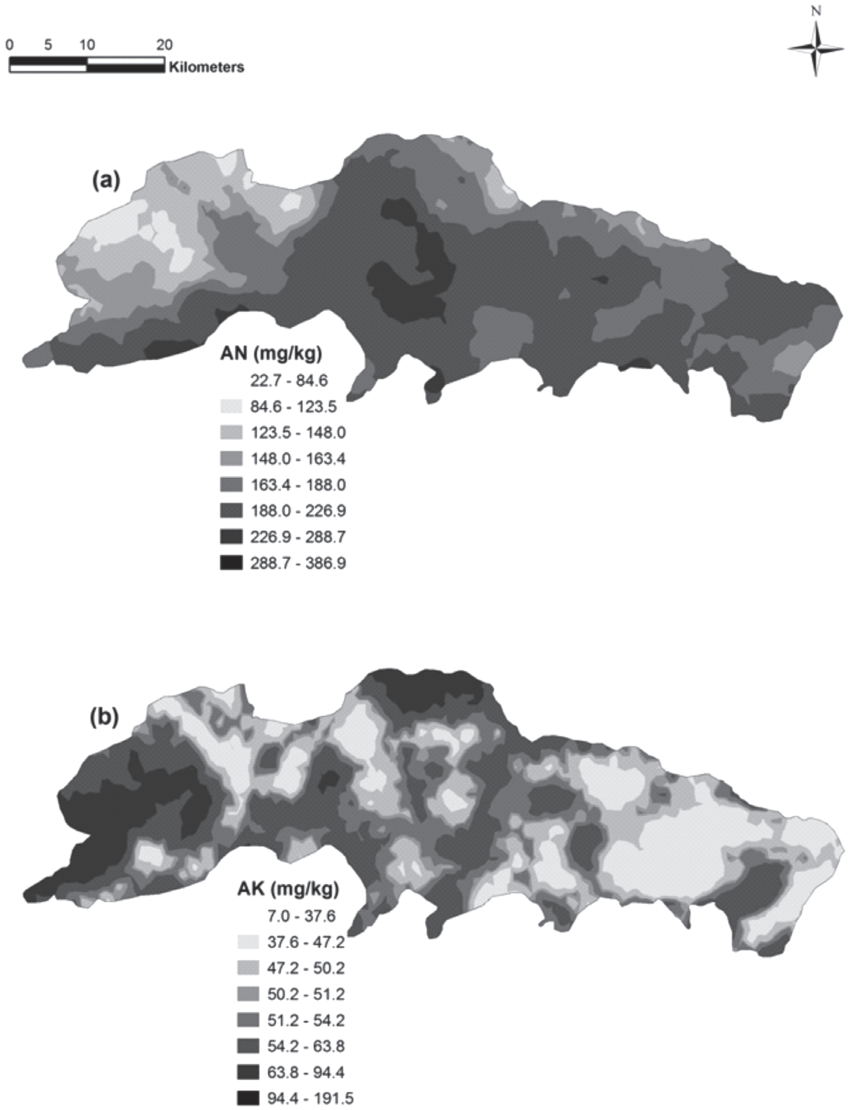

Fig. 3. Maps of the kriged estimates for (a) available nitrogen (AN) (mg kg^{-1}) and (b) available potassium (AK) (mg kg^{-1})

Evaluation indices resulting from cross-validation of spatial maps of soil properties are given in table 3. It was observed that**,** pH had low MAE and MSE; however, for OC, AN and AK relatively large MAE and MSE were observed. These results are in close conformity with the findings of Reza *et al.* (2010). For all the soil properties the G value was greater than 0, which indicates that spatial prediction using semivariogram parameters is better than assuming mean of observed value as the property value for any unsampled location. This also shows that semivariogram parameters obtained from fitting of experimental semivariogram values were reasonable to describe the spatial variation of pH, OC and AK. Because the RMSE value for AN was especially large, prediction of AN was especially poor suggesting that exponential model of kriging was unreliable for this parameter.

Table 3. Evaluation performance of kriged map of soil properties through cross-validation

Soil properties	MAE*	MSE**	G***
pH	0.005	0.229	99
Organic carbon	0.012	19.96	28
Available N	0.195	4129.5	5
Available K	0.362	445.0	7

*Mean absolute error; **Mean square error; ***Goodness of prediction

Relationship between Soil Properties

The Pearson linear correlation analysis results (Table 4) showed highly significant negative relation-

Table 4. Correlation coefficients among soil properties and their level of significance

Soil properties	pH	Organic carbon	Available N	Available K
pH	1.000			
Organic carbon	-0.434**	1.000		
Available N	-0.228**	0.498***	1.000	
Available K	0.194**	0.119**	0.063*	1.000

**Correlation significant at P=0.01; *Correlation significant at P=0.05

ships of soil pH with OC (r = -0.434**) and AN (r = -0.228**). There have been reports on a positive relationship between OC and the capacity of the soil to supply essential plant nutrients including N and K (Rezaei and Gilkes 2005). Pearson linear correlation analysis indicated significant positive relationships of OC with AN (r=0.498**) and AK (r = 0.119**). A positive correlation was also found between AN and AK (r = 0.063*).

Conclusions

The summary statistics for soil properties showed that there was difference in the CV of the soil properties. The raw data sets of pH and AK are strongly positively skewed and the application of log-transformation was effective in normalizing the data. Exponential model was best fitted with strongly spatially dependent and having spatial autocorrelations. The kriging interpolated map has shown that pH value in the western part of the study area was in neutral to alkaline range along with the Brahmputra river and the value in the southeast quadrant was strongly acidic. Organic carbon and AN had a large similar spatial variability. Both these soil properties decreased in the western part of the study area and increased in central, southwest and southeast quadrant. In general the geostatistical method on a large scale could be accurately used to evaluate spatial variability of soil properties in north-eastern India.

Acknowledgments

This research was funded by the Directorate of Agriculture, Government of Assam, under the project "Assessment and mapping of some important soil parameters for the thirteen priority districts of Assam state towards optimum land use planning".

References

Burgess, T.M. and Webster, R. (1980) Optimal interpolation and isarithmic mapping of soil properties. I. The semi-variogram and punctual kriging. *Journal of Soil Science* **31**, 315-331.

Cambardella, C.A., Moorman, T.B., Novak, J.M., Parkin, T.B., Karlen, D.L., Turco, R.F. and Konopka, A.E. (1994) Field-scale variability of soil properties in central Iowa soils. Soil *Science Society of America Journal* **58**, 1501-1511.

Chinchmalatpure, Anil R., Nayak, A.K. and Gururaja Rao, G. (2005) Spatial variability and mapping of extractable boron in soils of Bara tract under Sardar Sarovar Canal command area of Gujarat. *Journal of the Indian Society of Soil Science* **53**, 330-336.

Davis, B.M. (1987) Uses and abuses of cross-validation in geostatistics. *Mathematical Geology* **19**, 241-248.

ESRI (Environmental Systems Research Institute) (2009) ArcGIS 9.3.1 ESRI, Redlands, California.

Geypens, M., Vanongeval, L., Vogels, N. and Meykens, J. (1999) Spatial variability of agricultural soil fertility parameters in a Gleyic Podzol of Belgium. *Precision Agriculture* **1**, 319-326.

Gilbert, C. and Wayne, H. (2008) Kriging analysis of soil properties. *Journal of Soils and Sediments* **8**, 193-202.

Goovaerts, P. (1999) Geostatistics in soil science: State-of-the-art and perspectives. *Geoderma* **89**, 1-45.

Grewal, M.S., Dahiya, I.S. and Kuhad, M.S. (2001) Evaluation of spatial variation in available phosphorus content of soils of Haryana. *Journal of the Indian Society of Soil Science* **49**, 117-123.

Krige, D.G. (1951) A statistical approach to some basic mine valuation problems on the Witwatersrand. *Mining Society South Africa* **52**, 119-139.

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

Matheron, G. (1963) Principles of geostatistics. *Economic Geology* **58**, 1246-1266.

Mueller, T.G. and Pierce, F.J. (2003) Soil carbon maps: Enhancing spatial estimates with simple terrain attributes at multiple scales. *Soil Science Society of America Journal* **67**, 258-267.

Nayak, A.K., Chinchmalatpure, Anil R., Gururaja Rao, G., Khandelwal, M.K. and Tyagi, N.K. (2006) Spatial variability of DTPA extractable micronutrients in the soils of Bara tract of Sardar Sarovar Canal command in Gujarat state, India. *Journal of the Indian Society of Soil Science* **54**, 137-145.

Nayak, A.K., Chinchmalatpure, Anil R., Khandelwal, M.K., Gururaja Rao, G., Nath Abhay, Munshi, H.Y. and Tyagi, N.K. (2002) Geostatistical analysis of surface and sub-surface soil salinity of irrigation block of Sardar Sarovar Canal command. *Journal of the Indian Society of Soil Science* **50**, 283-287.

NBSS&LUP (1999) Soils of Assam for Optimizing Land Use. National Bureau of Soil Survey and Land Use Planning, Nagpur. *NBSS Publ. No.* **66.**

Santra, Priyabrata, Chopra, U.K. and Chakraborty, Debashis (2008) Spatial variability of soil properties and its application in predicting surface map of hydraulic parameters in an agricultural farm. *Current science* **95**, 937-945.

Quine, T.A. and Zhang, Y. (2002) An investigation of spatial variation in soil erosion, soil properties and crop production within an agricultural field in Devon U.K. *Journal of Soil and Water Conservation* **57**, 50-60.

Reza, S.K., Sarkar, D., Baruah, U. and Das, T.H. (2010) Evaluation and comparison of ordinary kriging and inverse distance weighting methods for prediction of spatial variability of some chemical parameters of Dhalai district, Tripura. *Agropedology* **20**, 38-48.

Rezaei, S.A. and Gilkes, R.J. (2005) The effects of landscape attributes and plant community on soil chemical properties in rangelands. *Geoderma* **125**, 167-176.

Schloeder, C.A., Zimmermen, N.E. and Jacobs, M.J. (2001) Comparison of methods for interpolating soil properties using limited data. *Soil Science Society of America Journal* **65**, 470-479.

Subbiah, B.V. and Asija, G.L. (1956) A rapid procedure for estimation of available nitrogen in soils. *Current Science* **25**, 259-260.

Sun, B., Zhou, S.L. and Zhao, Q.G. (2003) Evaluation of spatial and temporal changes of soil quality based on geostatistical analysis in the hill region of subtropical China. *Geoderma* **115**, 85-99.

Tsegaye, T. and Hill, R.L. (1998) Intensive tillage effects on spatial variability of soil test, plant growth, and nutrient uptake measurement. *Soil Science* **163**, 155-165.

Voltz, M. and Webster, R. (1990) A comparison of kriging, cubic splines and classification for redicting soil properties from sample information. *Journal of Soil Science* **41**, 473-490.

Walkley, A. and Black, I.A. (1934) An examination of the degtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. *Soil Science* **37**, 29-38.

Wang, Y., Xhang, X. and Huang, C. (2009) Spatial variability of soil total nitrogen and soil total phosphorus under different land uses in a small watershed on the loess plateau, China. *Geoderma* **150**, 141-149.

Warrick, A.W. (1998) Spatial variability. In *Environmental Soil Physics*. (D. Hillel, Ed.), pp. 655-675. Academic Press, New York.

Webster, R. and Oliver, M.A. (2001) *Geostatistics for Environmental Scientists*. John Wiley, Chichester, U.K.

Yao, L.X., Zhou, X.C., Cai, Y.F. and Chen, W.Z. (2004) Spatial variability of soil properties at different sampling intensities and accuracy of their estimation. *Soils* **36**, 538-542.

Yasrebi, J., Saffari, M., Fathi, H., Karimian, N., Emadi, M. and Baghernejad, M. (2008) Spatial variability of soil fertility properties for precision agriculture in Southern Iran. *Journal of Applied Sciences* **8**, 1642-1650.

Zhu, X.H., Yang, X.C. and Cai, Y.L. (2005) Fractal and fractal dimension of spatial distribution of China soil system. *Acta Pedologica Sinica* **42**, 881-888.

Received 14 July 2010; Received in revised form 24 January 2012; Accepted 21 May 2012