Agropedology 2019, 29 (01), 44-50

Developing Multiple-Linear Regression Model to Predict Soil Cation Exchange Capacity for Nagaland Soils

Gaurav Mishra*, Avishek Sarkar, Gopal Tiwari[1] and Abhishek Jangir[1]

Rain Forest Research Institute, Jorhat-785001, Assam, India
[1]ICAR-National Bureau of Soil Survey & Land Use Planning, Nagpur-440033, Maharashtra, India

Abstract: Cation exchange capacity (CEC) is an important parameter to assess the soil health and fertility. The procedure for measuring CEC is complicated and time consuming. To overcome this issue, researchers have developed and tested models to estimate CEC, but no such model has been developed for North-eastern region (NER) of India. In the present study, a training dataset of 198 numbers of soil samples having data of soil texture, bulk density (BD), pH, soil organic carbon (SOC) and CEC was used to develop step-wise regression model for CEC. Correlation analysis was done to extract the influential parameters for predicting CEC. Results showed that basic soil parameters were able to predict CEC and can define 90 % of variability with SSE value of 2.76. The agreement between observed and predicted CEC in validation dataset with R^2 value of 0.665 provided a strong basis to identify input variables for predicting CEC in the region.

Key words: *Cation exchange capacity, Soil organic carbon, Clay, Regression*

Introduction

Soil is the source of plant nutrients and various ecological services like water filtration, gas exchange, food supply and carbon storage (Lal *et al*. 2015; Khaledian *et al*. 2017). In addition to these ecosystem functions, soil is also equally important for sustaining the food production and human health (Keesstra *et al*. 2016 and Willaarts *et al*. 2016). Due to important contributions of soil, its health assessment is quite necessary which is prone to disturbance and/or land use change. The effect of disturbances can be assessed by analyzing change in its physical, biological and chemical parameters (Brejda *et al*. 2000). Askari and Holden, (2015) mentioned twenty-one soil properties, which can be used as potential indicators for soil health assessment. For determination of soil health, a huge data set containing a number of soil parameters is needed, which is not easily available and also not feasible to collect in the developing regions (Van Hall *et al*. 2017). Accordingly, the need of developing a cost effective and simple way to assess the soil quality is also advocated by several authors (Costa *et al*. 2015; Pulido *et al*. 2015).

Cation exchange capacity (CEC) has been used as an effective chemical indicator of soil quality in many studies (Masto *et al*. 2008 and Li *et al*. 2013a). CEC can be defined as the relative ability of a soil to hold and exchange cations (Saidi 2012). Brevik (2015) and Mukherjee and Zimmerman (2013) has already

* Corresponding author: (E-mail-gaurav.mishra215@gmail.com)

described the affect of soil physical (particle- size distribution), chemical (soil reaction) and biological properties (organic matter) on CEC. Even though the assessment of CEC is very important, we lack adequate datasets due to the classical and time consuming measurement (Budiman and Alfred 2011). To overcome this issue researchers have developed several pedotransfer functions (PTFs) to predict the CEC using statistical tools like multiple linear regressions (MLR) (Shabani and Norouzi 2015; Olorunfemi *et al*. 2016; Khaledian *et al*. 2016b; Sulieman *et al*. 2018). Researchers also advocated some important variables which can be used to predict CEC like soil organic matter (SOM) and clay content (Helling *et al*. 1964); pH, SOM and clay content (Seybold *et al*. 2005); inclusion of bulk density (Shekofteh *et al*. 2017) and recently SOM, pH, calcium carbonate equivalent (CCE) and soil texture (Khaledian *et al*. 2017).

As indicated in the literature, various types of PTFs were already developed using a variety of inputs and statistical approaches around the world, but no such attempt was made for the Northeast region (NER) of India. NER is characterized by unique agro-ecological conditions, making it the centre of speciation for many plant species (Choudhury *et al*. 2016). It is one of the twelve biodiversity hot-spots in the world with 65 % of its area under forests (Saha *et al*. 2012). However, the importance of CEC in evaluating the soil quality of forests is already mentioned Mishra *et al*. (2017). Therefore, this study aims to develop PTF for estimating CEC of Nagaland soils, using basic soil properties.

Materials and Methods

Study area

Nagaland state (25°10' to 27° 4' N; 93° 15' to 95° 15' E) in NER region of India cover an area of 1.66 million ha at an elevation ranging from 194 to 3,826 m above mean sea level (MSL) and falls under agro-ecological region AER 7**.** *The mean annual temperature (MAT) in the lower topography ranges from 23°to 24°C* whereas in the higher topography MAT varies from 16° to 17°C and the average annual rainfall exceeds more than 3000 mm. The majority of area in state is covered by *wet evergreen and moist deciduous* forests followed by agriculture land, mainly *Jhum lands* (Saha *et al*. 2012). Rice is the staple food of the state and occupies about 70 per cent of the total area under cultivation and constitutes about 75 per cent of the total food production in the State. Inceptisols and Entisols are the *major soils found in the area.*

Soil sampling and analys s

During the period of 5 years (2013 to 2018), 293 sites were randomly surveyed throughout the state and geo-referenced soil samples were collected from the surface layer (0-30 cm depth) with the help of GPS. Collected soil samples were immediately stored in polyethylene bags, air-dried and processed (2 mm sieved) to exclude litter, roots and coarse particles for laboratory analysis. For determination of percentage sand, silt and clay hydrometer method was used (Klute 1986). Core method was employed for estimation of Bulk density (BD). Soil pH was measured with 1:2.5 soil-water ratio. $K_2Cr_2O_7$ wet oxidation method (Walkley and Black1934) was used to estimate SOC content in soil. Cation exchange capacity (CEC) was estimated by 1 *N* ammonium acetate (pH 7.0) method (Sumner and Miller 1996).

Statistical analyses

Descriptive statistics such as minimum and maximum values, mean, standard deviation, and correlation (Pearson) were carried out in MS excel 2007. To remove the outliers, test of normality was performed and only 283 samples data set was retained for further analysis. The entire dataset (283 soil samples) were divided into two subsets, model development sub-set (70 %, 198) and validation data subset (30 %, 85). In order to develop CEC- PTF, stepwise multiple linear regression analysis was conducted in SPSS Version 16.0. After that, the efficiency of developed PTF was tested using different errors: 1) coefficient of determination (R^2), 2) adjacent coefficient of determination (Adj. R^2), and 3) standard error for the estimate (SEE).

Results and Discussion

Summary statistics of soil properties

Descriptive statistics for different soil parameters is provided in table 1. There was a huge gap between the minimum and maximum values of sand, silt and clay. The mean value of sand in Nagaland soils was 46.05 % (±13.64), with a maximum value of 90.75 %. The minimum value of silt and clay in dataset was 2.60 and 1.90 %. However, the mean value of silt and clay were 26.70 (±9.05) and 26.35 % (±10.64), respectively. After removing the outliers, the mean value of sand, silt and clay were 45.69 (±12.87), 26.52 (±8.76) and 27.59 % (±10.03), respectively. Among the textural components, sand percentage has highest mean value in dataset, followed by clay and silt. The variations in mean values of sand silt and clay along entire data set is mainly due to the land use management practices. Rao and Wagenet (1985) to suggested the impact of intrinsic (weathering) and anthropogenic (cultivation) factors in variation of basic soil parameters like soil texture. The minimum and maximum value of BD in dataset was 0.66 and 1.31 mg m^{-3}, with the mean value of 0.94 (±0.14) mg m^{-3}. However, after removing the outliers, mean value was 0.94 (±0.14) mg m^{-3}. The minimum and maximum value of BD in dataset was 0.66 and 1.31 mg m^{-3}, with the mean value of 0.94 (±0.13) mg m^{-3}. The variation in values of BD in the entire dataset is mainly due to the elevation gradient, as elevation ranges from 194 to 3,826 m above MSL in the state. Hanawalt and Whittaker (1976) also reported the negative relationship between BD and elevation. Similar kind of variation in BD with elevation was also reported by Mishra *et al*. (2018) in Nagaland state.

The soils under different land uses were acidic to neutral in nature. The minimum pH in dataset was 3.90, while the maximum value was 6.54. The mean value of soil pH in the data set was 5.10 (±0.74), while after removing the outliers, it was reduced to 5.04 (±0.70). The hilly terrain and high rainfall in the NER leads to the leaching of bases from exchange complex, which is the most important cause for acidic nature of soil in region. The minimum value of SOC in dataset was 0.47 %, while the mean value of SOC was 1.71 % (±0.65). Moreover, after removal of outlier, minimum and average values were changed to 0.47 and 1.79 % (±0.62), respectively. The differences recorded for SOC in soils of Nagaland is mainly due to the amount of litter inputs. Singh *et al*. (2014) also reported similar findings with high value of SOC in forest in comparison to other land uses from the same region. Moreover, the minimum and maximum value of CEC were 5.33 and 31.81 cmol (P^+) kg^{-1}. The mean value of CEC in Nagaland soils was 17.00 cmol (P^+) kg^{-1} (±6.58), but changed to 17.78 cmol (P^+) kg^{-1} (±6.61) after the removal of outliers. The variation in CEC can be supported with the findings of Brevik (2015) and Mukherjee and Zimmerman (2013), as they mentioned particle size distribution, pH and SOM the main drivers of CEC in soils.

Table 1. Descriptive statistics of measured soil properties in Nagaland soils.

Parameter	Sand (%)	Silt (%)	Clay (%)	BD (gm cc^{-3})	pH	SOC (%)	CEC (cmol kg^{-1})
Min	16.00	2.60	1.90	0.66	3.90	0.47	5.29
Max	90.75	58.00	57.15	1.31	6.54	3.01	31.81
Mean	46.05	26.70	26.35	0.94	5.10	1.71	17.00
Std. dev	13.64	9.05	10.64	0.14	0.74	0.65	6.58
Training dataset after test of normality (n=198)							
Min	16.00	6.15	3.45	0.68	3.93	0.50	5.33
Max	85.20	56.70	57.15	1.31	6.54	3.01	31.81
Mean	45.69	26.52	27.59	0.94	5.04	1.79	17.78
Std. dev	12.87	8.76	10.03	0.13	0.70	0.62	6.61

Correlation analysis

The correlation analysis between studied parameters and CEC are presented in table 2. It was found that silt (-0.575) and clay (-0.745) were negatively correlated with sand. There was significant negative correlation of CEC with sand (-0.153). Khaledian *et al.* (2017) also reported about the negative correlation between sand and CEC. However, clay (0.180) and SOC (0.924) were positively correlated with CEC in whole dataset. These results are in line with the findings of Ghorbani *et al*. (2015) who reported that clay and sand are better predictors of CEC in comparison to silt. Moreover, these relationships between particle-size and CEC suggest that the value of CEC in soil is more controlled by the mineralogy of parent material (Mousavi 2012). The relationship between SOC and CEC can be well explained by the findings of Brady and Weil (2008) who reported that SOC is strongly correlated with SOM and subsequently, SOC has the high value of CEC per unit volume (Khaledian *et al*. 2017). Zeraatpishe and Khormali (2012) also reported that high concentration of SOC is able to affect soil pH and therefore CEC also.

Table 2. Correlation matrix (Pearson) between different soil parameters and CEC.

Parameter	Sand	Silt	Clay	BD	pH	SOC	CEC
Sand	1.000						
Silt	-0.575**	1.000					
Clay	-0.745**	-0.116*	1.000				
BD	0.068	-0.051	-0.041	1.000			
pH	0.318**	-0.197**	-0.227**	0.198**	1.000		
SOC	-0.194**	0.047	0.198**	-0.088	-0.079	1.000	
CEC	-0.153**	0.007	0.180**	-0.102	-0.084	0.924**	1.000

**. Correlation is significant at the 0.01 level
*. Correlation is significant at the 0.05 level

Modeling CEC through stepwise regression

Step-wise regression analyses was performed on the training dataset, using sand, silt, clay, BD, pH and SOC to predict CEC and the results presented in table 3. The information about the regression standardized residual, through normal P-P plot is provided in figure 1. The results of step-wise regression analyses showed that all the studied parameters were included in generated PTF (Table 3). However, sand, clay and SOC were the parameters, which were significantly correlated to CEC in dataset (Table 2). The value of CEC is well predicted with studied parameters, with R^2 and SEE value of 0.904 and 2.76 (Table 3). Similar to methodology followed by Asadu *et al*. (1990) and Sulieman *et al*. (2018) step-wise regression model was developed for Nagaland soils using the different parameters. Scatter plots of the measured *versus* predicted CEC values for 85 soil samples from Nagaland state is shown in figure 2. The R^2 value for observed and predicted CEC, in the validation dataset was 0.665. It is reported that clay and SOC can define more than 50 % of the variation in CEC values. Moreover, role of SOC in controlling the CEC has already been explained (Ulusoy *et al*. 2016; Khaledian *et al*. 2016a). In our study, studied parameters are able to define 90 % (R^2 = 0.904) of the variability in CEC values of training dataset.

Table 3. Equation developed for soil cation exchange capacity (CEC) using step wise linear regression analysis.

Linear Regression models	R^2	Adjusted R^2	SSE
CEC = 14.621-0.128*sand-0.144*silt-0.132*clay+0.101*BD-0.092*pH+9.540*SOC	0.904	0.816	2.76

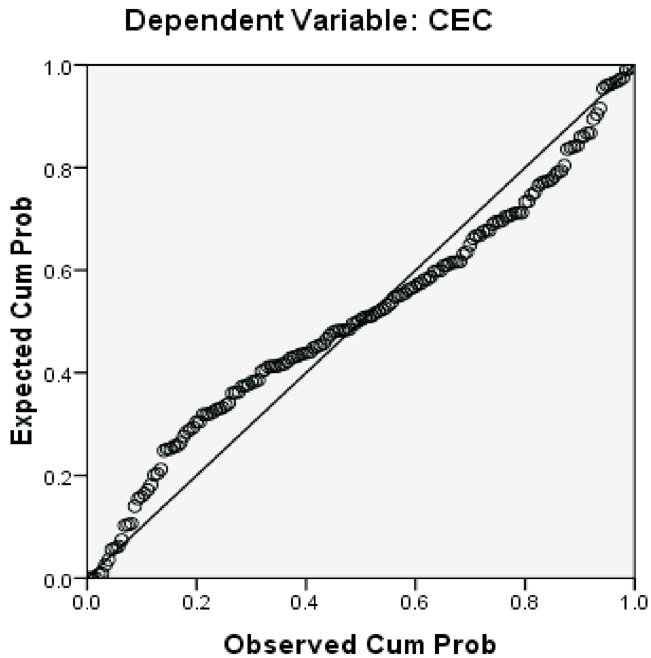

Fig. 1. Normal P-P Plot of regression standardized residual

Conclusion

Stepwise linear regression analysis was used to predict CEC. Sand, silt, clay, BD, pH and SOC were the basic inputs used to find their relationship with CEC in the data set. The correlation analysis showed that sand was negatively correlated, while clay and SOC were positively correlated with CEC. A stepwise linear regression model was developed using basic soil parameters to predict CEC and tested for its accuracy. The agreement between observed and predicted CEC values through different errors validated the findings that developed PTF can provide the precise estimate of CEC in the state/region concerned.

Acknowledgments

We thankfully acknowledge the financial support provided by Indian Council of Forestry Research and Education (ICFRE), Dehradun, India for this study. The authors declare that they have no conflict of interest.

References

Asadu, C.L.A. and Akamigbo, F.O.R. (1990). Relative contribution of organic matter and clay fractions to cation exchange capacity of soils in southeastern Nigeria. *Samaru Journal of Agricultural Research* **7,** 17-23.

Askari, M.S. and Holden, N.M. (2015). Quantitative soil quality indexing of temperate arable management systems. *Soil and Tillage Research* **150**, 57-67.

Brady, N.C. and Weil, R.R. (2008). 'Nature and Properties of Soils', 14th Edition. (Prentice Hall, Upper Saddle River, NJ).

Brejda, J.J., Moorman, T.B., Karlen, D.L., Smith, J.L. and Dao, T.H. (2000). Identification of regional soil quality factors and indicators: I. Central and southern hill plains. *Soil Science Society of America Journal* **64**, 2115-2124.

Brevik, E.C., and Sauer, T.J. (2015). The past, present, and future of soils and human health studies. *Soil* **1**, 35-46.

Budiman, M. and Alfred, E.H. (2011). Predicting soil properties in the tropics. *Earth Science Review* **106**, 52-62.

Choudhury, B.U., Fiyaz, A.R., Mohapatra, K.P. and Ngachan, S. (2016). Impact of land uses, agro-physical variables and altitudinal gradient on soil organic carbon concentration of North Eastern Himalayan Region of India. *Land Degradation and Development* **27**, 1163-1174.

Costa, J.L., Aparicio, V. and Cerdà, A. (2015). Soil physical quality changes under differentmanagement systems after 10 years in the Argentine humid pampa. *Solid Earth* **6**, 361-371.

Ghorbani, H., Kashi, H., Hafezi Moghadas, N. and Emamgholizadeh, S. (2015). Estimation of soil cation exchange capacity using multiple regression, artificial neural networks, and adaptive neuro-fuzzy inference system models in Golestan Province, Iran. *Soil Science and Plant Analysis* **46,** 763-780.

Hanawalt, R.B. and Whittaker, R.H. (1976). Attitudinally coordinated patterns of soils and vegetation in the San Jacinto Mountains, California. *Soil Science* **121,** 114-124.

Helling, C.S., Chesters, G. and Corey, R. (1964). Contribution of organic matter and clay to soil cation-exchange capacity as affected by the pH of the saturating solution. *Soil Science Society of America Journal* **28**, 517-520.

Keesstra, S.D., Bouma, J., Wallinga, J., Tittonell, P., Smith, P., Cerdà, A., Montanarella, L., Quinton, J.N., Pachepsky, Y., van der Putten, W.H., Bardgett, R.D., Moolenaar, S., Mol, G., Jansen, B. and Fresco, L.O. (2016). The significance of soils and soil science towards realization of the United Nations sustainable development goals. *Soil* **2**, 111-128.

Khaledian, Y., Kiani, F., Ebrahimi, S., Brevik, E.C. and Aitkenhead-Peterson, J. (2016a). Assessment and monitoring of soil degradation during land use change using multivariate analysis. *Land Degradation and Development*. (http://dx.doi.org/10.1002/ldr.2541).

Khaledian, Y., Pereira, P., Brevik, E.C., Pundyte, N. and Paliulis, D. (2016b). The influence of organic carbon and pH on heavy metals, potassium, and magnesium levels in Lithuanian Podzols. *Land Degradation and Development*. (http://dx.doi.org/10.1002/ldr.2638).

Khaledian, Y., Brevik, E.C., Pereirac, P., Cerdàd, A., Fattahe, M.A. and Tazikehf, H. (2017). Modeling soil cation exchange capacity in multiple countries. *Catena* **158,** 194-200.

Klute, A. (1986). Methods of Soil Analysis: Part 1. *In* 'Physical and Mineralogical Methods'. Soil Science Society of America Book Series No. 5. pp. 687-734 (Soil Science Society of America, Madison, Wisconsin.).

Lal, R., Negassa, W. and Lorenz, K. (2015). Carbon sequestration in soil. *Current Opinion Environmental Sustainability* **15,** 79-86.

Li, P., Zhang, T., Wang, X. and Yu, D. (2013a). Development of biological soil quality indicator system for sub-tropical China. *Soil and Tillage Research* **126**, 112-118.

Masto, R.E., Chhonkar, P.K., Singh, D. and Patra, A. K. (2008). Alternative soil quality indices for evaluating the effect of intensive cropping, fertilization and manuring for 31 years in the semi-arid soils of India. *Environmental Monitoring and Assessment* **136**, 419-435.

Mishra, G., Marzaioli, R., Giri, K., Borah, R., Dutta, A. and Jayaraj, R.S.C. (2017). Soil quality assessment under shifting cultivation and forests in Northeastern Himalaya of India. *Archives of*

Agronomy and Soil Sciences **63**, 1355-1368.

Mishra, G., Giri, K. and Pandey, S. (2018). Role of Alnus nepalensis in restoring soil Fertility: a case study in Mokukchung, Nagaland. *National Academy Science Letters*. (DOI. 10.1007/s40009-018-0668-4).

Mousavi, S.M. (2012). Application of Rosgen classification the Ziarat watershed, Gorgan. *Journal of American Science* **8,** 184-189.

Mukherjee, A. and Zimmerman, A.R. (2013). Organic carbon and nutrient release from a range of laboratory-produced biochars and biochar–soil mixtures. *Geoderma* **193-194**, 122-130.

Olorunfemi, E.I., Fasinmirin, T.J. and Ojo, S.A. (2016). Modeling cation exchange capacity and soil water holding capacity from basic soil properties. *European Journal of Soil Science* **5**, 266-274.

Pulido, M.M., Gabriels, D., Cornelis, W. and Lobo, D. (2015). Comparing aggregate stability tests for soil physical quality indicators. *Land Degradation and Development* **26,** 843-852.

Rao, P.S.C. and Wagenet, R.J. (1985). Spatial variability of field soils: methods for data analysis and consequences. *Weed Science* **33**, 18-24.

Saha, R., Chaudhary, R.S. and Somasundaram, J. (2012). Soil health management under hill agro-ecosystem of North East India. *Applied and Environmental Soil Science*. (https://doi.org/10.1155/2012/696174)

Saidi, D. (2012). Relationship between cation exchange capacity and the saline phase of Cheliff sol. *Agricultural Sciences* **3**, 434-443.

Seybold, C.A., Grossman, R.B. and Reinsch, T.G. (2005). Predicting cation exchange capacity for soil survey using linear models. *Soil Science Society of America Journal* **69**, 856-863.

Shabani, A. and Norouzi, M. (2015). Predicting cation exchange capacity by artificial neural network and multiple linear regression using terrain and soil characteristics. *Indian Journal of Science and Technology* 8**.** http://dx.doi.org/10.17485/ijst/2015/v8i28/83328.

Shekofteh, H., Ramazani, F. and Shirani, H. (2017). Optimal feature selection for predicting soil CEC: Comparing the hybrid of ant colony organization algorithm and adaptive network-based fuzzy system with multiple linear regressions. *Geoderma* **298**, 27-34.

Singh, A.K., Bordoloi, L.J., Kumar, M., Hazarika, S. and Parmar, B. (2014). Land use impact on soil quality in eastern Himalayan region of India. *Environmental Monitoring and Assessment* **186**, 2013-2024.

Sulieman, M., Saeed, I., Hassaballa, A. and Rodrigo Comino, J. (2018). Modeling cation exchange capacity in multi geo-chronological-derived alluvium soils: An approach based on soil depth intervals. *Catena*. (https://doi.org/10.1016/j.catena.2018.05.001).

Sumner, M.E. and Miller, W.P. (1996). Cation exchange capacity and exchange coefficients. *In* 'Methods of Soil Analysis Part 3, Chemical Methods'. (Eds. D.L. Sparks, , A. L. Page, P.A. Helmke,). pp. 1201-1229. (Soil Science Society of America, Madison, Wisconsin, USA).

Ulusoy, Y., Tekin, Y., Tümsavas, Z. and Mouazen, A.M. (2016). Prediction of soil cation exchange capacity using visible and near infrared spectroscopy. *Biosystems* **152**, 79-93.

Van Hall, R.L., Cammeraat, L.H., Keesstra, S.D. and Zorn, M. (2017). Impact of secondary vegetation succession on soil quality in a humid Mediterranean landscape. *Catena* **149,** 836-843.

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.

Willaarts, B.A., Oyonarte, C., Muñoz-Rojas, M., Ibáñez, J.J. and Aguilera, P.A. (2016). Environmental factors controlling soil organic carbon stocks in two contrasting Mediterranean-climate areas. *Land Degradation and Development* **27**, 603-611.

Zeraatpishe, M. and Khormali, F. (2012). Carbon stock and mineral factors controlling soil organic carbon in a climatic gradient, Golestan Province. ***Journal of Soil Science*** *and* ***Plant Nutrition*** **12,** 637-654.

Received: January, 2019 *Accepted: May, 2019*