

# Modelling soil cation exchange capacity in different land-use systems using artificial neural networks and multiple regression analysis

Gaurav Mishra<sup>1,\*</sup>, Juri Das<sup>1</sup> and Magboul Sulieman<sup>2,3</sup>

<sup>1</sup>Rain Forest Research Institute, Jorhat 785 001, India

<sup>2</sup>Soil Sciences Department, College of Food and Agricultural Sciences, King Saud University, P.O. Box 2460, Riyadh 11451, Saudi Arabia

<sup>3</sup>Department of Soil and Environment Sciences, Faculty of Agriculture, University of Khartoum, Khartoum North, 13314 Shambat, Sudan

Cation exchange capacity (CEC), as an important indicator of soil quality, represents the ability of the soil to hold positively charged ions. In this study, CEC was successfully predicted using different statistical methods, including artificial neural networks (ANNs) involving multi-layer perceptron (MLP), radial basis function (RBF), multiple linear regression (MLR) and nonlinear regression (NLR). About 293 soil samples were collected from North East India, which are under three land uses (shifting agriculture (jhum), forest and cash crops). Also, 70% of the samples (205 samples) was selected as the calibration set and the remaining 30% (88 samples) used as the prediction set. Soil pH, texture, bulk density (BD) and organic carbon (OC) were used as predictor variables to estimate CEC. The CEC-pedotransfer function (CEC-PTF) performance was evaluated with the coefficient of determination ( $R^2$ ), root mean square error (RMSE) and standard error for the estimate (SEE) between the observed and predicted values. The results indicated that the nonlinear model ( $R^2 = 0.91$  and  $SEE = 1.82$  for training) for cash-crop system, and RBF ( $R^2 = 0.91$  and  $SEE = 3.83$  for training) for jhum system were the best models to estimate CEC. In contrast, RBF ( $R^2 = 0.67$  and  $SEE = 14.87$  for training) for forest system was the worst model to estimate CEC. The results confirm that clay and OC were the most influential variables to predict CEC in the cash-crop system, whereas BD and OC were more suitable for jhum system. Although the ANNs provided suitable predictions of the entire dataset, NLR gave a formula to estimate soil CEC using commonly tested soil properties. Thus, NLR provided a reasonable estimate of CEC for most soils analysed.

**Keywords:** Artificial neural networks, cation exchange capacity, multiple regression, land uses.

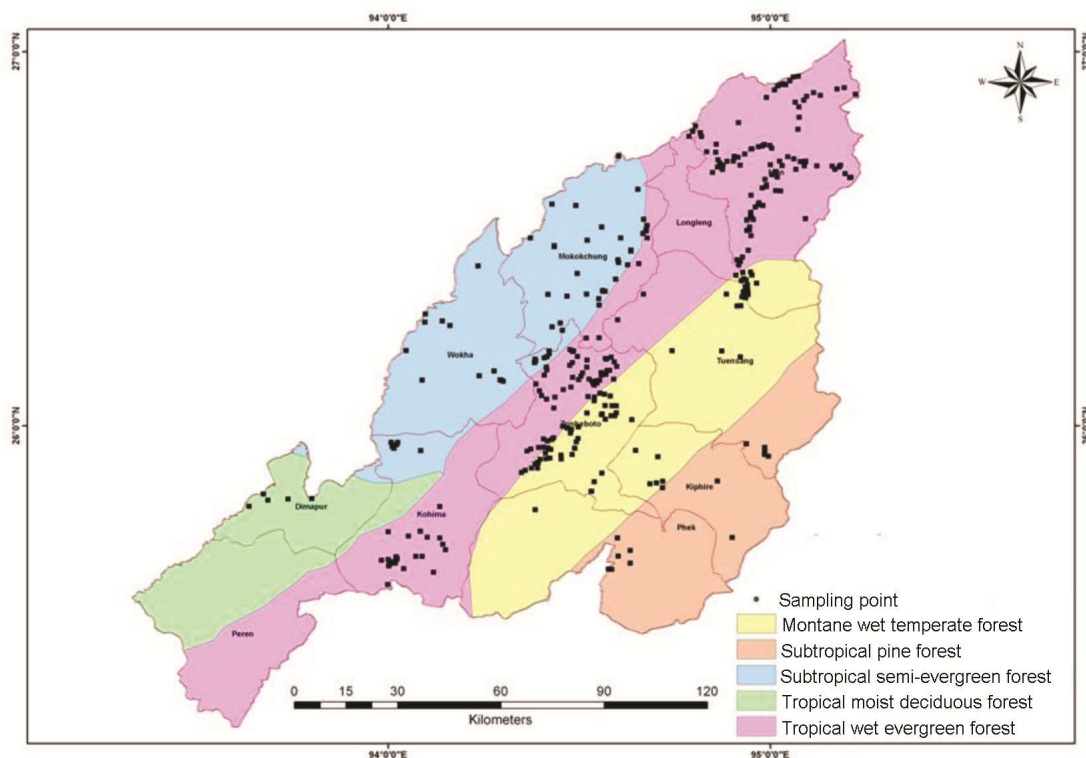
SOIL acts as a reservoir for different plant nutrients, and is well recognized to serve various ecological functions

like gas exchange, water filtration, food supply and carbon storage<sup>1-3</sup>. Apart from ecological services, soil also provides an important platform to various cycles like hydrological cycle, nutrient cycle and biological cycles, which are imperative for the sustainability of soil and human health<sup>4,5</sup>. Soil physical, biological and chemical parameters are the indicators of soil health and quality. However, these parameters are vulnerable to natural interruptions or land-use changes<sup>6</sup>. Therefore, a large dataset of various soil parameters is required for determination of soil quality. This task is generally not viable in developing regions due to less availability of resources<sup>7</sup>. Necessity of developing cost-effective and easy methods and ways for determination of soil quality has also been advocated by several researchers<sup>8,9</sup>.

Cation exchange capacity (CEC) is an important chemical property of soil that has been considered as a key indicator of soil quality in many studies<sup>10-12</sup>. It is the relative capacity of a soil to hold and exchange cations<sup>13</sup>. The effect of different soil physical (particle size distribution), chemical (soil reaction) and biological (organic matter) properties on CEC has been described in the literature<sup>14,15</sup>. Despite the significance of determination of CEC as an important indicator of soil chemical properties, adequate datasets are not available owing to its conventional and protracted measurement<sup>16-18</sup>. Researchers have developed different pedotransfer functions (PTFs) to predict CEC<sup>19-24</sup>. Seybold *et al.*<sup>25</sup> used soil organic matter (SOM), pH and clay content of soil as input soil properties to develop PTFs for CEC. Use of soil structural properties, like bulk density (BD), was also recommended by a few workers<sup>26,27</sup>. Khaledian<sup>3</sup> developed PTFs utilizing SOM, pH, calcium carbonate equivalent (CCE) and soil texture for different land uses of USA, Spain, Iran and Iraq.

As indicated in the above-mentioned studies, various types of PTFs have been developed using a variety of inputs and statistical approaches all over the world; however, no such attempt has been made for the North Eastern Region (NER) of India. Exceptional agro-ecological

\*For correspondence. (e-mail: gaurav.mishra215@gmail.com)



**Figure 1.** Soil sampling points for the entire dataset.

condition is one of the main characteristics of the NER of India, due to which it is known to be the centre of speciation for many plant species<sup>28</sup>. It is one among the 12 biodiversity hotspots in the world, with 65% of its area under forests and 16% under agriculture<sup>29</sup>. The shifting cultivation is a major agriculture practice popular among locals of this region and is one of the major sources of income for them. The importance of CEC in evaluating soil quality of forest and jhum land of this region has already been reported<sup>30</sup>. Thus, the present study aims to develop PTFs for estimating this important soil property using basic soil parameters in major land uses of NER of India.

## Materials and methods

### Study area

The present study was carried out in Nagaland. The state is located in the extreme northeast of the country, and lies between 25°10'–27°4'N lat. and 93°15'–95°15'E long., with an area of 1.66 M ha. The elevation ranges from 194 to 3826 m amsl. The average annual rainfall is 1831 mm, where 90% is distributed from May to October. The forest cover of the state is 12,489 sq. km (ref. 31). Temperate evergreen, tropical evergreen, tropical semi-evergreen, tropical moist deciduous, bamboo forests and degraded forests<sup>31</sup>. The major agricultural land-use pattern

is 'shifting cultivation', locally known as 'jhum'<sup>32</sup>. The upland rice (*Oryza sativa*) is the staple food of the state and farmers only use conventional practices for its cultivation<sup>33</sup>. Other major crops are maize, cowpea and vegetables. Tea and rubber are the major plantation crops of the state and are mainly cultivated in the foothill zones, especially in Mokukchung district. According to the soil taxonomy, most of soils in the region can be classified as Inceptisols followed by Ultisols<sup>34</sup>.

### Soil sampling and analysis

A total of 293 sites were randomly selected in the state, on the basis of land use (forest, jhum and cash crops) and elevation gradient. The sampling positions of all the sites were recorded with the help of GPS (Figure 1). Soil samples were collected from 0 to 30 cm depth in all the sites. They were immediately stored in polyethylene bags, after being air-dried at room temperature (22°C) and sieved through 2 mm sieve to remove litter, roots and coarse particles. Particle size analysis was done by the hydrometer method and percentage sand, silt and clay was calculated<sup>35</sup>. BD was estimated by the core method<sup>36</sup>. Soil pH was determined in 1:2 soil–water suspensions using a digital pH meter. Soil organic carbon (SOC) was estimated using the modified Walkley–Black method<sup>37</sup>. CEC was estimated by 1 N-ammonium acetate (pH 7.0) method<sup>38</sup>.

*Machine learning approaches*

**Artificial neural network:** The artificial neural network (ANN) is a powerful, soft computational technique which has been widely used in environmental sciences. It comprises of parallel systems that are composed of processing elements (PE) or neurons, which are assembled in layers and connected through several links or weights<sup>39</sup>. Multi-layer perception (MLP) and radial basis function (RBF) are the two most popular architectures used in ANN. MLP can be trained by a back-propagation algorithm, while RBF is multi-layer and feed-forward, and often used for strict interpolation in multi-dimensional space<sup>39</sup>. Before data analysis, the selected covariates were grouped into categories based on their land-use system. To establish PTFs, particular attention was given to determining covariates that were related to soil CEC. In order to determine the influential variables in different subsets, the correlation (Pearson) of variables was evaluated, and significant differences were estimated at  $P < 0.05$ . As the result, before obtaining the pedotransfer models, we ignored the covariates that did not have a significant effect on the estimation of soil CEC using a stepwise regression model.

*Statistical methods and data analysis*

Descriptive statistics such as minimum and maximum values, mean, standard deviation, variance, skewness and kurtosis was performed for each land use and the entire dataset. Linear correlation (Pearson) and multiple linear regression analyses were also conducted for all land uses and the whole dataset. After selecting the covariates that maximum influenced the prediction of soil CEC, we tested the fit of linear ( $y = ax + c$ ) and nonlinear quadratic regression ( $y = ax^2 + bx + c$ ) formulas for (i) the whole data regardless of land-use system, (ii) the forest system, and (iii) the shifting cultivation system and (iv) the cash crop system. We also optimized the models by ANNs, particularly RBF and MLP. Finally, the fitted graphs were prepared by testing 176 samples.

*Model validation*

All CEC-PTF models were evaluated based on four different error criteria, namely the coefficient of determination ( $R^2$ ), adjusted coefficient of determination ( $\text{adj.}R^2$ ), standard error for the estimate (SEE), and root mean square error (RMSE) between the observed and predicted values. These four indices were calculated using the following equations

$$R^2 = 1 - \frac{\sum_{i=1}^n (\text{pred}_i - \overline{\text{obs}})^2}{\sum_{i=1}^n (\text{obs}_i - \overline{\text{obs}})^2}, \quad (1)$$

$$\text{Adj } R^2 = 1 - (1 - R^2) \frac{n-1}{n - \text{pred}_i - 1}, \quad (2)$$

$$\text{SEE} = \sqrt{\frac{\sum (\text{obs}_i - \text{pred}_i)^2}{n}}, \quad (3)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\text{obs}_i - \text{pred}_i)^2}, \quad (4)$$

where  $n$  is the number of data points at the  $i$ th location;  $\text{obs}$  and  $\text{pred}$  are observed and predicted CEC values, and  $\overline{\text{obs}_i}$  denotes the mean for observed CEC values. SEE gives the bias, whereas RMSE estimates the prediction accuracy<sup>40</sup>. For all methods and datasets, about 70% of the data, was randomly selected to train the models and the remaining data (around 30%) was used as a validation dataset. All analyses were performed using SPSS 20 (IBM, USA).

**Results***Summary statistics of soil properties*

Table 1 presents a summary of the descriptive statistics for pH, sand content, silt content, clay content, BD, SOC and CEC. The pH was in the acidic range, with mean values ranging from 5.02 to 5.18. Also, not much variation was observed in the sand % (43.87–49.98 (m/m)) among land uses. The sand % was reported minimum in forests, while cash crops land-use system had the maximum value. The silt % varied from 26.46 to 27.88. Variation was also observed when the clay means were compared between different land uses (22.15%–29.67% (m/m)). The variation among sand, silt and clay is mainly related to the weathering, erosion, deposition and soil forming processes. The highest BD mean was recorded in the cash crop ( $1.06 \text{ kg m}^{-3}$ ) and lowest in jhum ( $0.87 \text{ kg m}^{-3}$ ). In general, BD was found to be low for all observed land uses. On the contrary, SOC content was high; it varied from 1.47% to 2.00% and highest SOC mean was observed in forest and lowest in jhum soils. CEC varied widely (13.60–20.19  $\text{cmol kg}^{-1}$ ), being highest in forest use and lowest in jhum soils.

*Correlation of CEC with other soil parameters*

Table 2 shows the Pearson's correlation results between CEC and sand, silt, clay, BD, pH and SOC. A positive correlation between CEC and sand ( $P < 0.05$ ) was observed in soils of forest and the whole systems. A non-significant relationship between silt and CEC ( $P < 0.05$ ) was recorded for the land uses except for forest. Our findings suggest a positive and significant correlation between

**Table 1.** Summary for descriptive statistics of measured soil properties under different land uses and whole dataset

Land-use system	Parameter	pH	Sand % (m/m)	Silt % (m/m)	Clay % (m/m)	Bulk density (kg m <sup>-3</sup> )	Soil organic carbon (SOC) (%)	Cation exchange capacity (CEC) (cmol kg <sup>-1</sup> )
Whole systems ( <i>n</i> = 293)	Minimum	3.61	6.00	2.00	0.35	0.56	0.31	4.35
	Maximum	6.82	94.60	73.6	81.80	1.45	3.21	34.42
	Mean	5.10	46.20	27.02	26.78	0.94	1.71	17.07
	Standard deviation	0.77	15.54	10.44	12.80	0.17	0.68	6.96
	Variance	0.60	241.42	109.0	163.8	0.03	0.46	48.51
	Skewness	0.36	0.47	0.97	1.01	0.36	0.06	0.23
	Kurtosis	-1.03	0.82	2.41	3.41	4.03	-0.92	-0.78
Forest ( <i>n</i> = 149)	Minimum	3.70	6.00	2.00	0.35	0.68	0.50	5.33
	Maximum	6.74	85.20	60.15	81.80	1.45	3.21	34.42
	Mean	5.07	43.87	26.46	29.67	0.97	2.00	20.19
	Standard deviation	0.79	13.88	9.61	13.18	0.15	0.60	6.41
	Variance	0.62	192.77	92.42	173.66	0.02	0.36	41.07
	Skewness	0.41	-0.084	0.04	1.56	1.98	-0.06	-0.06
	Kurtosis	-0.99	0.56	1.06	4.86	10.05	-0.86	-0.51
Jhum ( <i>n</i> = 107)	Minimum	3.61	12.85	3.70	1.50	0.56	0.31	4.99
	Maximum	6.54	94.60	73.60	57.80	1.19	2.78	25.39
	Mean	5.18	48.14	27.50	24.37	0.87	1.39	13.60
	Standard deviation	0.76	17.76	11.79	12.41	0.14	0.63	5.57
	Variance	0.58	315.39	139.10	154.02	0.02	0.40	30.98
	Skewness	0.30	0.83	0.90	-0.01	0.21	0.39	0.48
	Kurtosis	-1.25	0.53	2.93	-0.21	-0.64	-0.76	-0.80
Cash crops ( <i>n</i> = 37)	Minimum	3.71	24.00	2.60	7.20	0.81	0.37	4.35
	Maximum	6.82	78.00	58.00	40.85	1.36	2.59	29.98
	Mean	5.02	49.98	27.88	22.15	1.06	1.47	14.54
	Standard deviation	0.76	13.74	9.58	9.35	0.17	0.63	7.13
	Variance	0.57	188.77	91.84	87.39	0.03	0.40	50.79
	Skewness	0.44	-0.15	0.14	0.53	0.39	0.23	0.66
	Kurtosis	-0.42	-0.56	2.53	-0.68	-1.16	-1.30	-0.49

**Table 2.** Correlation (Pearson) between soil covariates and cation exchange capacity (CEC) for different land-use systems

Land-use system	Sand	Silt	Clay	BD	pH	SOC
Whole systems	-0.153**	0.007ns	0.18**	-0.102ns	-0.084ns	0.924**
Forest	-0.163*	0.166*	0.052ns	-0.324**	-0.07ns	0.903**
Jhum	0.043ns	-0.108ns	0.042ns	0.19*	-0.034ns	0.942**
Cash crops	-0.277ns	-0.038ns	0.445**	0.026ns	-0.155ns	0.865**

Significant differences are indicated as  $P < 0.05^*$  and  $P < 0.01^{**}$ . ns, Not significant.

CEC and clay content ( $P < 0.05$ ) for the cash crops and the whole systems. Significant positive relationship was established between CEC and BD ( $P < 0.05$ ) under forest and jhum land use. The pH of the studied soils showed a non-significant relationship with CEC. On the contrary, SOC of all the studied soils showed a positive and highly significant ( $P < 0.05$ ) relationship with CEC.

#### Estimation of CEC using regression analysis

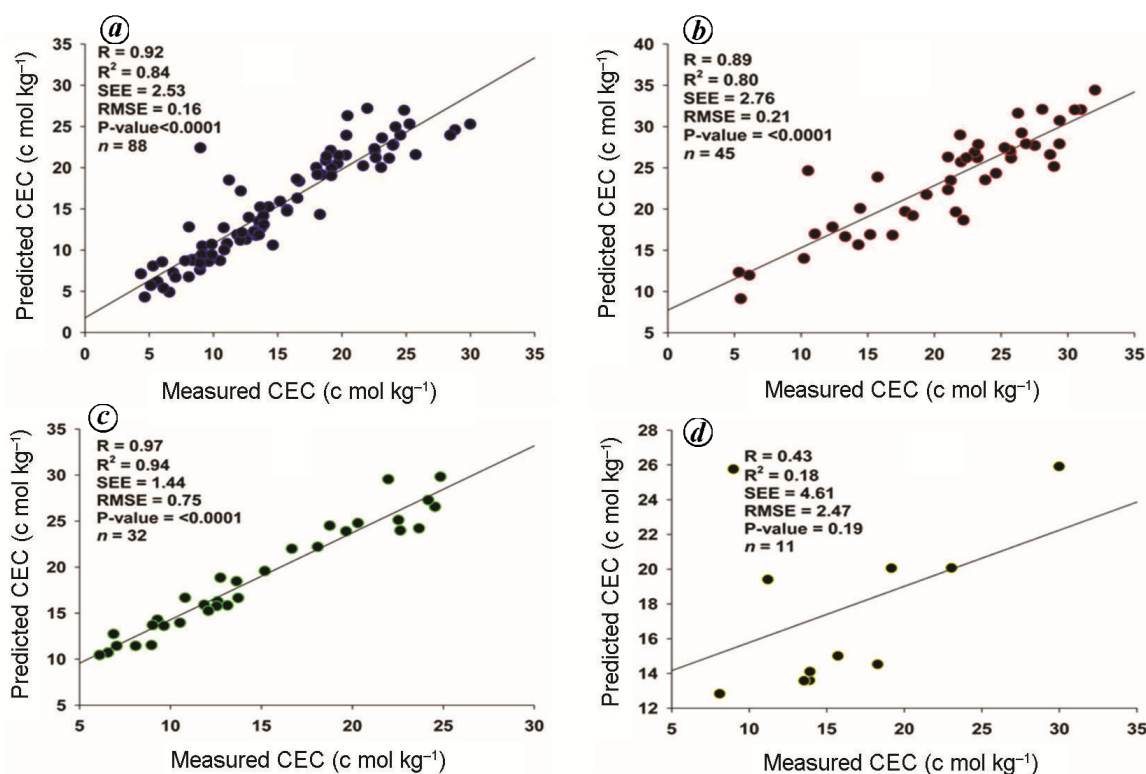
Table 3 shows the results of multiple linear regression analyses for CEC, under the whole systems and different land uses. In the whole systems, CEC is well predicted by SOC when linear regression method was used.  $R^2$  and

SSE values for the equation are 0.86 and 2.66 respectively. Moreover, when nonlinear method was used,  $R^2$  and SSE values remained similar. In forest soils, CEC was well predicted with SOC, with  $R^2$  and SSE values of 0.80 and 2.76 respectively (Figure 2). Although  $R^2$  and SSE values were the same under both the linear and nonlinear regression, the number of covariates was more in the nonlinear regression equation. Similar results for both the equations were obtained for jhum lands also. The value for  $R^2$  under jhum was 0.89, which is better than forests. However, in plantation crops, the values of  $R^2$  and SSE for linear and nonlinear regression equations were different (Figure 3). For linear equation the values were 0.75 and 3.63, while for nonlinear equation they were 0.92 and 1.82 respectively.

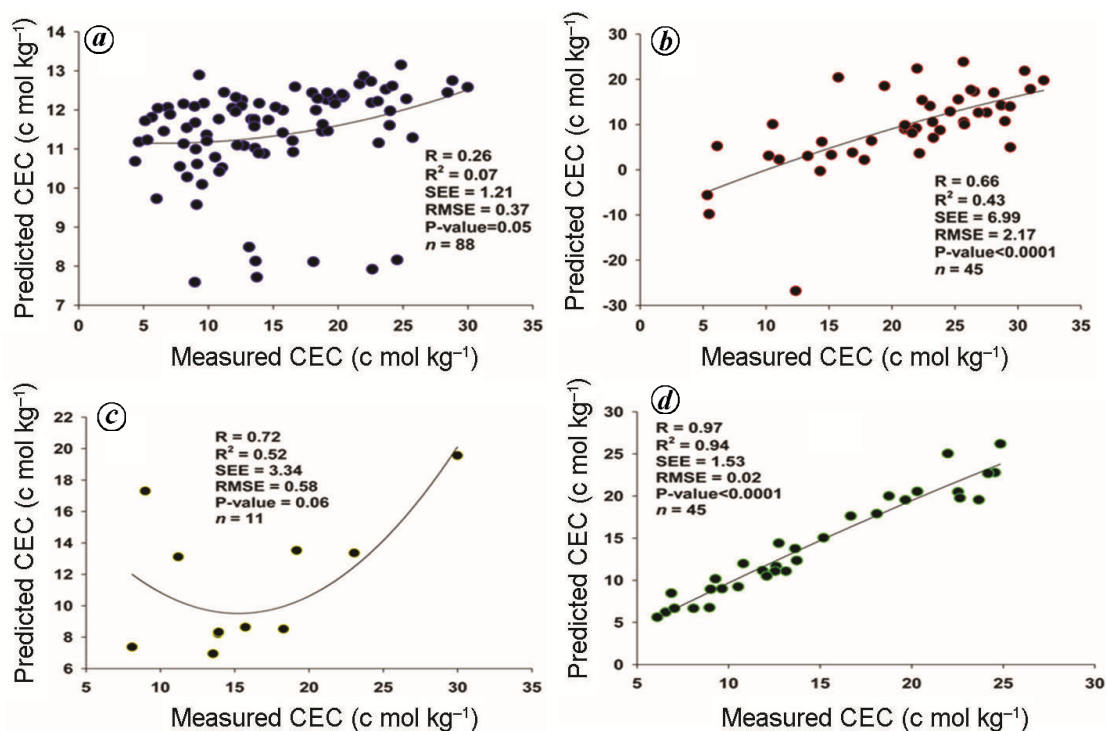
**Table 3.** Performance indices for the training models with their equations

Land-use system	Training set	Regression method	R	R <sup>2</sup>	SEE	P-value	Equation
Whole systems	205	Linear <sup>a</sup>	0.93	0.86	2.66	0.00	CEC = 0.025*Sa + 0.021*Cl + 9.506*SOC - 0.858
		Linear <sup>b</sup>	0.93	0.86	2.66	0.00	CEC = 9.652*SOC + 0.724
		Non-linear <sup>a</sup>		0.86	2.66	0.00	CEC = -0.246*Sa <sup>2</sup> + 0.00*Cl <sup>2</sup> + 0.086*SOC <sup>2</sup> + 0.073*Sa + 0.036*Cl + 0.086*SOC + 9.359
		Non-linear <sup>b</sup>		0.86	2.66	0.00	CEC = 0.149*SOC <sup>2</sup> + 9.117*SOC + 1137
Forest	104	Linear <sup>a</sup>	0.90	0.82	2.77	0.00	CEC = 0.013*Sa + 0.021*Si - 1.415*BD + 9.516*SOC + 1.47
		Linear <sup>b</sup>	0.90	0.82	2.76	0.00	CEC = 9.626*SOC + 1.027
		Non-linear <sup>a</sup>		0.84	2.76	0.00	CEC = 0.001*Sa <sup>2</sup> + 0.001*Si <sup>2</sup> - 28.127*BD <sup>2</sup> + 0.188*SOC <sup>2</sup> - 0.133*Sa - 0.46*Si + 55.329*BD + 8.659*SOC - 20.053
		Non-linear <sup>b</sup>		0.82	2.76	0.00	CEC = -0.474*SOC <sup>2</sup> + 11.496*SOC - 0.648
Jhum	75	Linear <sup>a</sup>	0.94	0.89	1.86	0.00	CEC = 8.233*SOC - 2.017*BD + 8.394
		Linear <sup>b</sup>	0.94	0.89	1.86	0.00	CEC = 8.301*SOC + 2.051
		Non-linear <sup>a</sup>		0.86	1.86	0.00	CEC = 7.901*BD <sup>2</sup> - 0.164*SOC <sup>2</sup> - 16.574*BD + 8.662*SOC + 10.075
		Non-linear <sup>b</sup>		0.89	1.86	0.00	CEC = -0.183*SOC <sup>2</sup> + 8.856*SOC + 1.707
Cash crops	26	Linear <sup>a</sup>	0.87	0.76	3.60	0.00	CEC = 0.09*Cl + 9.228*SOC - 1.056
		Linear <sup>b</sup>	0.87	0.75	3.63	0.00	CEC = 9.762*SOC + 0.152
		Non-linear <sup>a</sup>		0.92	1.82	0.00	CEC = 0.005*Cl <sup>2</sup> + 2.299*SOC <sup>2</sup> - 0.225*Cl + 0.155*SOC + 5.974
		Non-linear <sup>b</sup>		0.75	3.61	0.00	CEC = 1.327*SOC <sup>2</sup> + 5.671*SOC + 2.785

For regression method: <sup>a</sup>Regression model using enter method for all covariates; <sup>b</sup>Regression model using stepwise method. SEE, Standard error of the estimate.



**Figure 2.** Predicted versus measured values of soil cation exchange capacity (CEC) of validation set using linear regression model with performance indices for: *a*, whole land-use systems; *b*, forest; *c*, jhum; *d*, cash crops.



**Figure 3.** Predicted versus measured values of soil CEC of validation set using nonlinear regression (quadratic) model with performance indices for: *a*, whole land-use systems; *b*, forest; *c*, jhum; *d*, cash crops.

### Estimation of CEC using ANNs

Table 4 shows the results of RBF and MLP functions. The  $R^2$  value for whole system under RBF network was 0.86 and the number of hidden layers was 9, while in jhum lands the value of  $R^2$  was the highest (0.91) and 8 hidden layers were found. The number of input and hidden layer was similar under MLP network of the whole systems and jhum lands, but the latter had slightly better values of  $R^2$  (0.89) compared to the former.

### Discussion

The pH was found to be acidic in nature, indicative of no salinity problem in the studied area. The NER of India is mainly a hilly region endowed with heavy rainfall, which leads to the leaching of bases from the exchange complex. Our findings suggest that the textural composition of the soil were inter-related to each other, where the decrease in one may increase the other. The variation in mean values of sand, silt and clay in the land uses is mainly due to management practices. This suggests formation of clay from sand and silt fractions due to alteration and/or neof ormation under humid subtropical condition<sup>41</sup>. Rao and Wagenet<sup>42</sup> suggested that variation in basic soil parameters like soil texture is due to intrinsic (weathering) and anthropogenic (cultivation) factors. BD

varied widely, being highest in the cash crop and lowest in jhum land use. This variation in BD can be explained by the differences in organic matter content, cultivation process and biotic activities<sup>43</sup>. SOC content was in general high in the studied soils. Jhum lands had the highest values of SOC, due to continuous slash down and burning of plant litter and its easy decomposition. Low CEC was observed in the studied soils which may be due to the presence of low-activity clay (kaolinite), as CEC is immensely affected by mineralogy of the soil<sup>44</sup>. It was found that CEC was least in jhum due to continuous washing away of topsoil because of steep slopes and high rainfall in the studied zone. In all the studied land uses, direct and positive correlation between CEC and SOC was noticed. This showed that CEC was directly affected by organic matter<sup>45</sup>. Singh *et al.*<sup>32</sup> also reported high value of SOC in forest compared to other land uses from the same region. The variation in CEC for the different land uses is also supported by the results Brevik<sup>14</sup>, and Mukherjee and Zimmerman<sup>15</sup>, who reported that particle size distribution, pH and SOM are the main drivers of CEC in soils.

In all the land uses and the entire dataset, significant positive relationship between SOC and CEC was recorded. This is in accordance with the findings of Brady and Weil<sup>46</sup>, who reported that SOC is strongly correlated with SOM and subsequently, SOC has a high value of CEC per unit volume<sup>47</sup>. Zeraatpishe and Khormali<sup>48</sup> reported that

**Table 4.** Summary of the radial basis function (RBF) and multi-layer perceptron (MLP) networks and their performance indexes used in this study

Land-use system	Artificial neural networks method	Training set			Validation set			$R^2$ (quadratic) of observed versus predicted	Input layer	Hidden layer	Output layer
		TN	SSE	RE	TN	SEE	RE				
Whole systems	RBF	213	12.12	0.11	80	8.20	0.20	0.86	1	9	1
	MLP	202	15.49	0.15	91	6.35	0.16	0.85	1	1	1
Forest	RBF	99	14.87	0.30	50	8.78	0.40	0.67	4	2	1
	MLP	106	8.79	0.47	43	4.75	0.20	0.83	4	3	1
Jhum	RBF	79	3.83	0.10	28	1.24	0.08	0.91	1	8	1
	MLP	73	3.73	0.10	34	1.58	0.13	0.89	1	1	1
Cash crops	RBF	25	2.11	0.18	12	3.15	0.36	0.71	2	5	1
	MLP	26	3.72	0.30	11	0.91	0.43	0.79	2	3	1

TN, Total number of samples; SEE, Sum of squares error; RE, Relative error.

high concentration of SOC affects soil pH and thereby CEC. SOC is able to explain maximum variation in CEC under different land uses and different techniques. Under jhum lands, 89% of variation has been explained by SOC by linear regression technique. However, in total dataset, SOC can explain 86% of variability in CEC. The role of SOC in controlling CEC has been already explained by several researchers<sup>16,48,49</sup>. Our results are also in line with their findings, as SOC is the basic input which can be significantly utilized to predict CEC in this region.

## Conclusion

In this study, ANNs were employed to predict CEC. Sand, silt, clay, BD, pH and SOC were the basic inputs used to find their relationship with CEC among different land uses and the entire dataset. Correlation analysis showed that SOC was the important property having relationship with CEC in land uses and entire dataset. The most influential variables for the cash crop system were clay and SOC, whereas BD and SOC showed the most influenced on jhum system. Overall, 16 different models were developed using different techniques and tested for different land uses and the entire dataset for their accuracy. Although the ANNs provided suitable predictions of the entire dataset, nonlinear regression provided a formula to estimate soil CEC using commonly tested soil properties.

- Brevik, E. C. and Sauer, T. J., The past, present, and future of soils and human health studies. *Soil*, 2015, **1**, 35–46.
- Lal, R., Negassa, W. and Lorenz, K., Carbon sequestration in soil. *Curr. Opin. Environ. Sustain.*, 2015, **15**, 79–86.
- Khaledian, Y., Brevik, E. C., Pereirac, P., Cerdà, A., Fattahe, M. A. and Tazikehf, H., Modeling soil cation exchange capacity in multiple countries. *Catena*, 2017, **158**, 194–200.
- Willaarts, B. A., Oyonarte, C., Muñoz-Rojas, M., Ibáñez, J. J. and Aguilera, P. A., Environmental factors controlling soil organic carbon stocks in two contrasting Mediterranean-climate areas. *Land Degrad. Dev.*, 2016, **27**, 603–611.
- Keesstra, S. D. *et al.*, The significance of soils and soil science towards realization of the United Nations sustainable development goals. *Soil*, 2016, **2**, 111–128.
- Brejda, J. J., Moorman, T. B., Karlen, D. L., Smith, J. L. and Dao, T. H., Identification of regional soil quality factors and indicators: I. Central and southern hill plains. *Soil Sci. Soc. Am. J.*, 2000, **64**, 2115–2124.
- Van Hall, R. L., Cammeraat, L. H., Keesstra, S. D. and Zorn, M., Impact of secondary vegetation succession on soil quality in a humid Mediterranean landscape. *Catena*, 2017, **149**, 836–843.
- Costa, J. L., Aparicio, V. and Cerdà, A., Soil physical quality changes under different management systems after 10 years in the Argentine humid pampa. *Solid Earth*, 2015, **6**(1), 361–371.
- Pulido, M. M., Gabriels, D., Cornelis, W. and Lobo, D., Comparing aggregate stability tests for soil physical quality indicators. *Land Degrad. Dev.*, 2015, **26**(8), 843–852; <http://dx.doi.org/10.1002/ldr.2225>.
- Larson, W. E. and Pierce, F. J., The dynamics of soil quality as a measure of sustainable management. In *Defining Soil Quality for a Sustainable Environment* (ed. Doran, J. W.), SSSA Spec. Pub. No. 35, ASA, CSSA and SSSA, Madison, WI, 1994, pp. 37–51.
- Li, P., Zhang, T., Wang, X. and Yu, D., Development of biological soil quality indicator system for subtropical China. *Soil Tillage Res.*, 2013, **126**, 112–118.
- Masto, R. E., Chhonkar, P. K., Singh, D. and Patra, A. K., Alternative soil quality indices for evaluating the effect of intensive cropping, fertilization and manuring for 31 years in the semi-arid soils of India. *Environ. Monit. Assess.*, 2008, **136**(1–3), 419–435.
- Saidi, D., Relationship between cation exchange capacity and the saline phase of Cheliff sol. *Agric. Sci.*, 2012, **3**(3), 434–443; <https://doi.org/10.4236/as.2012.33051>.
- Brevik, E. C., Soil health and productivity. In *Soils, Plant Growth and Crop Production* (ed. Verheye, W.), Encyclopedia of Life Support Systems (EOLSS), EOLSS Publishers, Oxford, UK, developed under the auspices of UNESCO; <http://www.eolss.net>
- Mukherjee, A. and Zimmerman, A. R., Organic carbon and nutrient release from a range of laboratory-produced biochars and biochar–soil mixtures. *Geoderma*, 2013, **193–194**, 122–130.
- McBratney, A. B., Minasny, B., Cattle, S. R. and Vervoort, R. W., From pedotransfer function to soil inference system. *Geoderma*, 2002, **109**, 41–73.
- Amini, M., Abbaspour, K. C., Khademi, H., Fathianpour, N., Afyuni, M. and Schulin, R., Neural network models to predict cation exchange capacity in arid regions of Iran. *Eur. J. Soil Sci.*, 2002, **53**, 748–757.
- Budiman, M. and Alfred, E. H., Predicting soil properties in the tropics. *Earth Sci. Rev.*, 2011, **106**, 52–62.
- Bell, M. A. and Van Keulen, H., Soil pedotransfer functions for four Mexican soils. *Soil Sci. Soc. Am. J.*, 1995, **59**, 865–871.
- Yukselen, Y. and Kaya, A., Prediction of cation exchange capacity from soil index properties. *Clay Miner.*, 2006, **41**, 827–837.

21. Shabani, A. and Norouzi, M., Predicting cation exchange capacity by artificial neural network and multiple linear regression using terrain and soil characteristics. *Indian J. Sci. Technol.*, 2015, **8**(28), 1–10; <http://dx.doi.org/10.17485/ijst/2015/v8i28/83328>.
22. Olorunfemi, E. I., Fasinmirin, T. J. and Ojo, S. A., Modeling cation exchange capacity and soil water holding capacity from basic soil properties. *Eur. J. Soil Sci.*, 2016, **5**(4), 266–274.
23. Khaledian, Y., Pereira, P., Brevik, E. C., Pundyte, N. and Paliulis, D., The influence of organic carbon and pH on heavy metals, potassium, and magnesium levels in Lithuanian Podzols. *Land Degrad. Dev.*, 2016, **28**(1), 345–354; <http://dx.doi.org/10.1002/ldr.2638>.
24. Sulieman, M., Saeed, I., Hassaballa, A. and Rodrigo Comino, J., Modeling cation exchange capacity in multi geochronological-derived alluvium soils: an approach based on soil depth intervals. *Catena*, 2018, **167**, 327–339; <https://doi.org/10.1016/j.catena.2018.05.001>.
25. Seybold, C. A., Grossman, R. B. and Reinsch, T. G., Predicting cation exchange capacity for soil survey using linear models. *Soil Sci. Soc. Am. J.*, 2005, **69**, 856–863.
26. Hartmann, A., Gräsele, W. and Horn, R., Cation exchange processes in structured soils at various hydraulic properties. *Soil Tillage Res.*, 1998, **47**(1), 67–72.
27. Shekofteh, H., Ramazani, F. and Shirani, H., Optimal feature selection for predicting soil CEC: comparing the hybrid of ant colony organization algorithm and adaptive network-based fuzzy system with multiple linear regression. *Geoderma*, 2017, **298**, 27–34.
28. Choudhury, B. U., Fiyaz, A. R., Mohapatra, K. P. and Ngachan, S., Impact of land uses, agrophysical variables and altitudinal gradient on soil organic carbon concentration of North Eastern Himalayan Region of India. *Land Degrad. Dev.*, 2016, **27**(4), 1163–1174.
29. Saha, R., Chaudhary, R. S. and Somasundaram, J., Soil health management under hill agroecosystem of North East India. *Appl. Environ. Soil Sci.*, 2012, **2012**, 1–9.
30. Mishra, G., Marzaioli, R., Giri, K., Borah, R., Dutta, A. and Jayaraj, R. S. C., Soil quality assessment under shifting cultivation and forests in Northeastern Himalaya of India. *Arch. Agron. Soil Sci.*, 2017, **63**(10), 1355–1368.
31. FSI, Indian's state of forest report. Forest Survey of India, Dehradun, 2017, p. 363.
32. Singh, A. K., Bordoloi, L. J., Kumar, M., Hazarika, S. and Parmar, B., Land use impact on soil quality in eastern Himalayan region of India. *Environ. Monit. Assess.*, 2014, **186**, 2013–2024.
33. Rathore, S. S., Paradigm shift for enhancing rice productivity in Nagaland: existing practices and their refinement. *Himalayan Ecol.*, 2008, **16**(2), 17–25.
34. Soil Survey Staff, *Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys*, Natural Resources Conservation Service. US Department of Agriculture Handbook 436, 1999, 2nd edn.
35. Klute, A. (ed.), *Methods of Soil Analysis: Part I. Physical and Mineralogical Methods*, Soil Science Society of America (SSSA), Book Series No. 5, Madison, Wisconsin, USA, 1986, pp. 687–734.
36. Blake, G. R. and Hartge, K. H., In *Methods of Soil Analysis Part I, Physical and Mineralogical Methods* (ed. Klute, A.), SSSA Book Series No. 5. Soil Science Society of America, Madison, Wisconsin, USA, 1986, 2nd edn, pp. 363–375.
37. Walkley, A. and Black, I. A., An examination of the Degtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. *Soil Sci.*, 1934, **37**(1), 29–38.
38. Sumner, M. E. and Miller, W. P., Cation exchange capacity and exchange coefficients. In *Methods of Soil Analysis Part 3, Chemical Methods* (eds Sparks, D. L., Page, A. L. and Helmke, P. A.), SSSA, Madison, Wisconsin, USA, 1996, pp. 1201–1229.
39. Memarian, H. and Balasundram, S. K., Comparison between Multi-layer perceptron and radial basis function networks for sediment load estimation in a tropical watershed. *J. Water Resour. Protec.*, 2012, **4**, 870–876.
40. Verfaillie, E., Lancker, V. V. and Meirvenne, M. V., Multivariate geostatistics for the predictive modeling of the surficial sand distribution in shelf seas. *Cont. Shelf Res.*, 2006, **26**, 2454–2468.
41. Karmakar, R. M. and Rao, A. E. V., Soils on different physiographic units in Lower Brahmaputra Valley zone of Assam I. Characterization and classification. *J. Indian Soc. Soil Sci.*, 1999, **47**, 761–767.
42. Rao, P. S. C. and Wagenet, R. J., Spatial variability of field soils: methods for data analysis and consequences. *Weed Sci.*, 1985, **33**, 18–24.
43. Rao, V. A. P., Naidu, M. V. S., Ramavatharam, N. and Rama Rao, G., Characterization, classification and evaluation of soils on different landforms in Ramachandrapuram mandal of Chittoor district in Andhra Pradesh for sustainable land use planning. *J. Indian Soc. Soil Sci.*, 2008, **56**, 23–33.
44. Bhattacharyya, T., Sen, T. K., Singh, R. S., Nayak, D. C. and Sehgal, J. L., Morphology and classification of Ultisols with Kandic horizon in north eastern region. *J. Indian Soc. Soil Sci.*, 1994, **42**, 301–306.
45. Nayak, D. C., Gangopadhyay, S. K. Sarkar, D. K. and Sen, T. K., Characteristics and classification of some acid soils of lower Subansiri district of Arunachal Pradesh. *Agropedology*, 2002, **12**, 112–121.
46. Brady, N. C. and Weil, R. R., *Nature and Properties of Soils*, Prentice Hall, NJ, USA, 2008, 14th edn.
47. Zeraatpishe, M. and Khormali, F., Carbon stock and mineral factors controlling soil organic carbon in a climatic gradient, Golestan Province. *J. Soil Sci. Plant Nutr.*, 2012, **12**(4), 637–654.
48. Ulusoy, Y., Tekin, Y., Tümsavas, Z. and Mouazen, A. M., Prediction of soil cation exchange capacity using visible and near infrared spectroscopy. *Biosyst. Eng.*, 2016, **152**, 79–93; <http://dx.doi.org/10.1016/j.biosystemseng.2016.03.005>.
49. Khaledian, Y., Kiani, F., Ebrahimi, S., Brevik, E. C. and Aitkenhead-Peterson, J., Assessment and monitoring of soil degradation during land use change using multivariate analysis. *Land Degrad. Dev.*, 2016, **28**(1), 128–141; <http://dx.doi.org/10.1002/ldr.2541>.

ACKNOWLEDGEMENTS. We thank the Indian Council of Forestry Research and Education, Dehradun, for providing financial support. The authors declare that they have no conflict of interest.

Received 24 December 2018; revised accepted 13 March 2019

doi: 10.18520/cs/v116/i12/2020-2027