Modeling soil cation exchange capacity in multiple countries

Yones Khaledian[a], Eric C. Brevik[b,*], Paulo Pereira[c], Artemi Cerdà[d], Mohammed A. Fattah[e], Hossein Tazikeh[f]

[a] Department of Agronomy, Iowa State University, Ames, IA, USA
[b] Department of Natural Sciences, Dickinson State University, Dickinson, ND, USA
[c] Environmental Management Center, Mykolas Romeris University, Ateities g. 20, LT-08303 Vilnius, Lithuania
[d] Soil Erosion and Degradation Research Group, Department of Geography, University of Valencia, Valencia, Spain
[e] Department of Soil and Water Science, Faculty of Agricultural Sciences, University of Sulaimani, Iraq
[f] Department of Soil Sciences, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran

ARTICLE INFO

ABSTRACT

Cation exchange capacity (CEC), as an important indicator for soil quality, represents soil's ability to hold positively charged ions. We attempted to predict CEC using different statistical methods including monotone analysis of variance (MONANOVA), artificial neural networks (ANNs), principal components regressions (PCR), and particle swarm optimization (PSO) in order to compare the utility of these approaches and identify the best predictor. We analyzed 170 soil samples from four different nations (USA, Spain, Iran and Iraq) under three land uses (agriculture, pasture, and forest). Seventy percent of the samples (120 samples) were selected as the calibration set and the remaining 50 samples (30%) were used as the prediction set. The results indicated that the MONANOVA ($R^2 = 0.82$ and Root Mean Squared Error (RMSE) = 6.32) and ANNs ($R^2 = 0.82$ and RMSE = 5.53) were the best models to estimate CEC, PSO ($R^2 = 0.80$ and RMSE = 5.54) and PCR ($R^2 = 0.70$ and RMSE = 6.48) also worked well and the results were very similar to each other. While the most influential variables for the various countries and land uses were different and CEC was affected by different variables in different situations, clay (positively correlated) and sand (negatively correlated) were the most influential variables for predicting CEC for the entire data set. Although the MANOVA and ANNs provided good predictions of the entire dataset, PSO gives a formula to estimate soil CEC using commonly tested soil properties. Therefore, PSO shows promise as a technique to estimate soil CEC. Establishing effective pedotransfer functions to predict CEC would be productive where there are limitations of time and money, and other commonly analyzed soil properties are available.

1. Introduction

Soils are an important source of services (filtration of water, exchange of gases with the atmosphere, foundation material for construction) and resources (food, fiber, fuel, clay for construction, medications) for humans (Mol and Keesstra, 2012; Brevik and Sauer, 2015). Soils play a pivotal role in the Earth system to manage the hydrological, erosional, geochemical, and biological cycles (Smith et al., 2015; Willaarts et al., 2016) which makes sustainable soil management and conservation necessary to achieve a sustainable world (Keesstra et al., 2016). Soil conservation includes the maintenance of soil fertility and quality. To determine soil quality, we need a large amount of data that is difficult to obtain due to a lack of measurements, particularly in developing regions, because of their high cost (Masto et al., 2015; Van

Leeuwen et al., 2015; Van Hall et al., 2017). The development of easy, accurate and low cost ways to determine soil properties important in the evaluation of soil quality to improve the management of the soil system in areas where data is scarce or non-existent is needed (Costa et al., 2015; Pulido et al., 2015).

Cation exchange capacity (CEC) is a measure of the ability of a soil to hold and exchange cations (Saidi, 2012). Determining soil quality requires identification of soil properties that are important in a soil's ability to carry out its various functions as well as being responsive to changes in land use and land management (Paz-Ferreiro and Fu, 2016; Zolfaghari et al., 2016). CEC plays an important role in soil quality (Brevik, 2009; Khaledian et al., 2016a; Taghizadeh-Mehrjardi, 2016). CEC can be influenced by soil physical (e.g., soil texture), chemical (e.g., pH, mineralogy), and biological (e.g., soil organic matter)

---

characteristics (Mukherjee and Zimmerman, 2013); therefore, CEC integrates aspects of all three of the indicator groups commonly used to determine soil quality (Brevik, 2009) and there are strong positive relationships between CEC and the soil colloidal fraction (Kweon et al., 2013). However, the conventional analytical methods used for measuring CEC such as the sodium saturation method (Chapman, 1965) are time and money consuming (Busenberg and Clemency, 1973; Seybold et al., 2005; Yukselen and Kaya, 2008).

Pedotransfer functions (PTF) are developed to use data from soil surveys or other sources to predict soil properties that are more difficult to determine in terms of cost and/or time and are therefore determined less frequently than more easily determined properties such as texture, pH, and organic carbon (OC) (McDonald, 1998; Borggaard et al., 2004; Reidy et al., 2016). These functions are useful to help scientists estimate properties like CEC in order to predict the spatial distribution of these important soil properties. Pedotransfer functions have often been developed using multiple linear regression models (Cornelis et al., 2001; Sequeira et al., 2014).

There are a number of studies on statistical methods to estimate soil parameters and assist in environmental management (i.e., McBratney and Odeh, 1997; Chen et al., 2014; Khaledian et al., 2016a; Hosseini et al., 2017) including CEC using general linear models (Seybold et al., 2005), multiple linear regression (Shabani and Norouzi, 2015; Khaledian et al., 2016b), PLS and stepwise regression (Khaledian et al., 2016a), adaptive network-based fuzzy inference system (ANFIS) and artificial neural networks (ANNs) (Ghorbani et al., 2015), and genetic expression programming (GEP) and multivariate adaptive regression splines (MARS) (Emamgolizadeh et al., 2015). However, these studies have not modeled CEC using advanced statistical methods such as algorithms (particle swarm optimization (PSO) and monotone regression tool (MONANOVA)) and compared the performance of these methods, which would represent an advance in the current knowledge.

In spite of the fact that estimation of ecosystem changes using statistical methods is a priority for soil and environmentalist scientists, selecting the correct and most productive methods is still an area that needs additional study. Hence, the goals of this study were to estimate soil CEC with pedo-transfer functions developed using intelligent (PSO, ANNs, and MONANOVA) and regression (PCR) methods, and compare these methods in order to find the best and most productive model(s). In doing so, we could determine the most effective soil properties for modeling CEC in the various countries and land uses evaluated.

2. Materials and methods

2.1. Study areas and soil sampling

Soil samples were collected from three different land uses: agriculture, pasture, and forest lands, in four countries (the USA, Spain, Iran and Iraq) with soil samples taken to a depth of 100 cm. Therefore, large ranges in soil chemical and physical properties were expected.

2.2. Physical and chemical analysis

Soil CEC was determined using the sodium acetate (NaOAc) method (Chapman, 1965). OC was determined using a wet combustion method (Nelson and Sommers, 1982). Soil pH was determined using a pH meter (model WTW 7110) and a 1:1 water/soil suspension (McLean, 1982). Calcium carbonate equivalent (CCE) was measured by the titration method (Page et al., 1982). Soil texture was determined by the Bouyoucos hydrometer method (Gee and Bauder, 1986).

2.3. Statistical analysis

Descriptive analyses such as average, minimum, maximum, standard deviation, and correlation (Pearson) were carried out using SPSS Version 16.0 (IBM Corporation, Armonk, NY). The data (170 soil samples) were divided into a calibration data subset (70%, 120) and prediction data subset (30%, 50). In order to model the data in various countries and land uses, a leave-one-out full cross-validation (LOOV) was used for validation of the models.

Data subsets were used for determining the performance of MONANOVA, PSO and PCR. Matlab (MathWorks, Natick, MA) was used to analyze PSO and XLSTAT (Addinsoft, Paris, France) was used for MONANOVA and PCR.

2.4. ANNS, MONANOVA, PCR, and PSO

Artificial neural networks (ANNs) are a computational model that is part of machine learning. ANNs depend on a large collection of associated artificial neurons similar to axons in a biological brain. In fact, the objective of ANNs is to solve the issues in the same way that human brains do. ANNs focus on not only biological processes in the brain but also neural networks' application to artificial intelligence (McCulloch and Pitts, 1943). In this study, the network was designed with 10 and 1 nodes in the input and output layers, respectively. The minimum possible difference between real and estimated data was aimed in various layers during the training process.

MONANOVA is the combination of a monotonic transformation of responses to a linear regression, like OLS (ordinal linear regression), as a method to improve the linear regression results. MONANOVA works based on the ALS (alternating least squares) algorithm. Its principle consists of alternating between a conventional estimation, such as linear regression or ANOVA, and a monotonic transformation of the dependent variables. The MONANOVA algorithm was introduced by Kruskal (1965).

Particle swarm optimization (PSO) was invented from research into simulating social behavior as an evolutionary technique which makes few or no assumption regarding the issue optimizing and can seek in a very large space. PSO optimizes a problem iteratively, improving a solution by considering a given measure of quality. It solves an issue by a population of particles, or candidate solutions, that these particles move around in the search-space based on simple mathematical formula in different positions and velocities. Each particle's movement is affected by its local best known position and updated as better positions are found by the other particles. Therefore, swarms move toward the best solutions (Kennedy and Eberhart, 1995).

PCR is a regression analysis technique that works based on principal component analysis (PCA). PCR regresses the response (the outcome) on a set of covariates (as predictors) based on a standard linear regression model; however, it uses PCA for predicting the unknown regression coefficients. Furthermore, one of the major uses of PCR is to overcome the multicollinearity problem (Jolliffe, 1982).

To find the best model to estimate soil CEC, the accuracy indicators R^2 and RMSE were used. High coefficient of determination values (R^2) and low RMSE values demonstrate that the models are capable.

3. Results

Descriptive statistics regarding CEC in the study areas are shown in Table 1. The lowest mean CEC values were found in Spain and the highest mean values were found in the agricultural soils of Iran (38.89 cmol/kg) and Iraq (36.7 cmol/kg), followed by CEC values in forest lands (37.55 cmol/kg) in Iran (Table 1). The pasture soils investigated in the USA and Iran had similar CEC values, 26.56 and 27.92 cmol/kg, respectively (Table 1). In terms of the other soil characteristics in different land uses and countries, respectively, the highest levels of clay and silt were measured in forest and agricultural soils, respectively, and Iran, whereas the highest percentage of sand was found in pastures in the USA and Spain. The mean CCE in agricultural and Spanish soils was twice that in the other land uses and countries. Pasture soils had the highest amount of OC. The variation in pH was not remarkable, as in all lands and countries the value was around 7 (Table 1).

Table 1
Descriptive statistics for CEC in the different land uses and countries covered by this study. *N* = number of samples.

County	Land use	N	Indicators	CEC (cmol/kg)	OC (%)	pH	CCE (%)	Clay (%)	Silt (%)	Sand (%)
Iran	Agriculture	8	Minimum	20.5	0.5	7.3	5	35	38.33	4
			Maximum	58	3.5	7.7	20	45	54	26.66
			Mean	39.89	2.1	7.54	10.94	40.5	47.21	12.42
			Std. deviation	15.09	1.13	0.13	6.11	3.68	6.19	7.81
	Pasture	20	Minimum	10	0.1	6.6	3	10	13	10
			Maximum	45	7	7.5	74	49	56.66	77
			Mean	27.92	2.54	7.26	17.87	37.2	38.9	24.15
			Std. deviation	8.38	1.6	0.26	17.84	10.22	9.43	17.47
	Forest	37	Minimum	13	0.2	5	1.5	22	11	3
			Maximum	75	8.7	8.2	72	69	60	67
			Mean	37.55	2.44	6.88	14.2	44.56	36.86	19.12
			Std. deviation	12.75	2.22	0.77	16.41	11.76	10.74	15.19
Iraq	Agriculture	40	Minimum	17.44	0.46	6.91	1	8.38	30.16	1.66
			Maximum	61.25	3.19	7.9	69.5	62.31	75.66	45.26
			Mean	36.7	1.67	7.52	18.74	40.48	47.2	12.31
			Std. deviation	10.84	0.73	0.26	15.24	14.27	9.19	10.2
Spain	Agriculture	40	Minimum	6.34	0.81	7.7	34	18.98	36.66	29.7
			Maximum	12.53	5.82	8.7	59	27.23	45.54	42.11
			Mean	8.91	2.39	8.315	46.25	22.47	41.94	35.58
			Std. deviation	1.63	1.43	0.21	6.64	2.09	2.21	3.56
USA	Pasture	25	Minimum	13.4	1.5	6.2	5.2	14.6	22.1	18.9
			Maximum	41.2	55.3	8	7.3	35.4	49.6	63.3
			Mean	26.56	16.23	7.18	6.4	26.06	37.07	36.87
			Std. deviation	7.09	14.66	0.62	0.78	5.78	7.11	12.14
All countries	All soils	170	Minimum	6.34	0.1	5	1	8.38	11	1.66
			Maximum	75	55.3	8.7	74	69	75.66	77
			Mean	27.98	4.27	7.48	21.94	34.62	41.25	24.27
			Std. deviation	14.76	7.57	0.69	18.9	13.13	9.04	15.29

3.1. Correlation analysis

Table 2 shows the correlations between the studied variables in the entire data set (Table 2a) as well as in the various land uses (Table 2b) and countries (Table 2c) considered in the study. In the different lands and countries (except for Spain), CEC was positively influenced by clay and negatively influenced by sand, with all of the clay correlations being > 0.67. CEC was significantly and negatively correlated with pH in agricultural land uses in Spain, whereas in the other categories, there was no significant correlation with pH. Even though there was a significant positive relationship between CEC and OC in Spain, the USA, Iraq, and pasture, CEC was not correlated with OC in the entire data set, Iran, agricultural soils, or forest soils.

3.2. Estimation of soil CEC in various land uses

The results of MONANOVA ($R^2 = 0.943$ and RMSE = 4.06) and PCR ($R^2 = 0.884$ and RMSE = 5.812) (Table 3) best described variation in agricultural lands, followed by pasture and forest lands, respectively (Table 3). Soil CEC was most influenced by sand and clay in agricultural and forest areas, and sand in pasture areas. The accuracy indicators (high R^2 and low RMSE) showed that MONANOVA was more powerful than PCR and PSO in estimating soil CEC.

3.3. Estimation of soil CEC for various countries

As can be seen in Table 3, MONANOVA was able to predict soil CEC better than PCR for all land uses (Table 3). The highest R^2 values for MONANOVA were obtained in the USA and Iraq, and RMSE values were also low. The lowest RMSE value was obtained for Spain (Table 3). The lowest R^2 and highest RMSE values were obtained for samples from Iran. Therefore, estimation of CEC in the USA and Iraq were the most promising, with CEC estimation in Spain also showing good promise. Even the Iran R^2 value was fairly good. CEC values were primarily influenced by clay and sand in Iran, clay and silt in Iraq, OC in Spain, and sand, silt and OC in the USA.

3.4. Estimation of soil CEC for the entire data set

CEC was positively influenced by clay and negatively by sand and pH. OC and silt did not show significant relationships with CEC (Table 2). MONANOVA ($R^2 = 0.822$ and RMSE = 6.32), ANNs ($R^2 = 0.82$ and RMSE = 5.53) and PSO ($R^2 = 0.86$ and RMSE = 5.543) had more promising results than PCR ($R^2 = 0.707$ and RMSE = 6.48) (Table 3).

4. Discussion

Research on CEC has been mainly descriptive, with CEC used as a parameter to characterize soils (Behera and Shukla, 2015; Mahmoud and Ghoneim, 2016). Soil CEC is highly affected by soil physical (e.g., soil texture), chemical (e.g., pH, mineralogy), and biological (e.g., OC) properties (Khaledian et al., 2016a); consequently, there is a strong and significant relationship between CEC and soil properties that are commonly analyzed. For example, Shabani and Norouzi (2015) observed that CEC had a significant negative correlation with sand content. However, this relationship is not always true. Amini et al. (2005) found that sand and silt did not have a significant effect in predicting CEC. Ghorbani et al. (2015) concluded that silt has less impact on the estimation of soil CEC than clay and sand. High amounts of OC can affect soil pH and, therefore, CEC (Zeraatpishe and Khormali, 2012). In addition, SOM, which is strongly correlated to OC, has a high CEC per unit volume (Brady and Weil, 2008). In general, soil CEC has not been well investigated using novel and advanced statistical approaches (Khaledian et al., 2016a); therefore, further work in this area is essential.

The statistically positive and negative relationships observed in this study (with the exception of the Spanish soils) between CEC with clay and sand, respectively, is supported by other studies (McBratney et al., 2002; Ulusoy et al., 2016). In contrast, CEC was primarily associated with OC content in Spain, as has been seen in other studies (Khaledian et al., 2012; Ulery et al., 2017). Although Seybold et al. (2005) and Khaledian et al. (2016a) observed that CEC increased with OC content,

Table 2
The correlation matrix (Pearson) between variables in the entire data set, various countries and land uses. Number of samples (N). Significant differences were considered at a $p < 0.05$.

(a)

The entire data set	$N = 170$					
Variables	Clay	Silt	Sand	pH	OC	CEC
Clay	1					
Silt	−0.067	1				
Sand	−0.813	−0.518	1			
pH	−0.406	0.105	0.289	1		
OC	−0.120	0.022	0.086	−0.112	1	
CEC	0.839	−0.008	−0.712	−0.471	0.072	1

(b)

Forest	$N = 37$						Pasture	$N = 45$				
Variables	Clay	Silt	Sand	pH	OC	CEC	Clay	Silt	Sand	pH	OC	CEC
Clay	1						1					
Silt	−0.026	1					0.570	1				
Sand	−0.751	−0.608	1				−0.906	−0.862	1			
pH	−0.151	−0.289	0.359	1			−0.014	0.113	−0.046	1		
OC	−0.227	−0.035	0.167	−0.358	1		−0.098	0.369	−0.133	0.085	1	
CEC	0.752	−0.090	−0.526	0.103	−0.247	1	0.669	0.658	−0.746	−0.106	0.443	1

Agriculture	$N = 88$					
Variable	Clay	Silt	Sand	pH	OC	
Clay						
Silt	−0.174	1				
Sand	−0.865	−0.343	1			
pH	−0.592	−0.305	0.720	1		
OC	−0.066	−0.187	0.157	0.155	1	
CEC	0.883	0.100	−0.893	−0.762	−0.086	1

(c)

Iran	$N = 65$						Iraq	$N = 40$				
Variables	Clay	Silt	Sand	pH	OC	CEC	Clay	Silt	Sand	pH	OC	CEC
Clay	1						1					
Silt	0.087	1					−0.702	1				
Sand	−0.767	−0.686	1				−0.766	0.080	1			
pH	−0.239	−0.121	0.273	1			−0.079	0.042	0.072	1		
OC	−0.166	−0.033	0.121	−0.250	1		0.536	−0.312	−0.469	0.211	1	
CEC	0.693	0.116	−0.572	−0.023	−0.189	1	0.946	−0.730	−0.666	−0.042	0.528	1

Spain	$N = 40$						USA	$N = 25$				
Variables	Clay	Silt	Sand	pH	OC	CEC	Clay	Silt	Sand	pH	OC	CEC
Clay							1					
Silt	0.369	1					0.772	1				
Sand	−0.817	−0.838	1				−0.928	−0.953	1			
pH	0.632	0.201	−0.496	1			0.059	0.253	−0.176	1		
OC	−0.357	0.249	0.055	−0.540	1		0.565	0.786	−0.730	0.154	1	
CEC	−0.191	0.318	−0.085	−0.551	0.869	1	0.826	0.878	−0.907	0.039	0.859	1

Units: CEC (cmol/kg), OC, CCE, clay, silt and sand (%).

different results were observed in this study in the agriculture, forest and Iran categories. These results indicated that soil CEC was more strongly controlled by clay than soil OC, as described by Zeraatpishe and Khormali (2012). The different behavior of the soils in Spain might due to the changes in soil properties triggered by long-term intensive agriculture, which changed the soil properties due to ploughing and the use of pesticides. Agricultural soils in many parts of Spain have been found to be highly degraded, characterized by surface crusting and high soil erosion rates (Taguas et al., 2015; Rodrigo-Comino et al., 2016; Martínez-Casasnovas et al., 2016). The soils in Spain were also enriched in calcium carbonate due to their limestone parent materials and OC levels were not high; therefore, CEC was explained mostly by pH. The carbon exhaustion of the soils in the Mediterranean, such as was the case with the Spanish soils, is due to the fact that the Mediterranean

Table 3
Performance indices (R² and RMSE) and the equations developed by the models tested.

Models	Categories	No.	R²	Adjusted R²	RMSE	Formula
MONANOVA	The entire date set	170 (50 for testing)	0.822	0.817	6.32	
	Agriculture	88	0.943	0.94	4.06	
	Forest	37	0.723	0.679	7.21	
	Pasture	45	0.883	0.868	2.78	
	Iran	65	0.72	0.695	7.01	
	Iraq	40	0.973	0.97	1.87	
	Spain	40	0.905	0.894	0.52	
	USA	25	0.97	0.964	1.36	
PCR	The entire date set	170	0.707	0.676	6.48	CEC = 27.842 + 0.706*Clay − 0.092*Silt − 0.169*Sand − 2.277*pH + 0.278*OC
	Agriculture	88	0.884	0.877	5.812	CEC = 41.025 + 1.091*Clay + 0.512*Silt + 0.159*Sand − 10.087*pH + 0.494*OC
	Forest	37	0.623	0.57	8.417	CEC = 10.973 + 0.427*Clay − 0.379*Silt − 0.436*Sand + 4.313*pH + 6.963E-02*OC
	Pasture	45	0.745	0.713	4.094	CEC = −64.880 + 1.365*Clay + 0.992*Silt + 0.864*Sand − 2.453*pH + 0.282*OC
	Iran	65	0.512	0.47	9.186	CEC = 9.471 + 0.475*Clay − 0.236*Silt − 0.346*Sand + 3.125*pH − 0.233*OC
	Iraq	40	0.906	0.895	3.517	CEC = 39.964 + 0.339*Clay − 0.453*Silt − 0.296*Sand + 0.964*pH + 0.475*OC
	Spain	40	0.82	0.80	0.731	CEC = 25.219 + 0.132*Clay + 2.167E-02*Silt − 0.054*Sand − 2.445*pH + 0.865*OC
	USA	25	0.932	0.914	2.090	CEC = −2408.584 + 24.647*Clay + 24.442*Silt + 24.212*Sand − 1.392*pH + 0.243*OC
PSO	The entire date set	170 (50 for testing)	0.80	0.82	5.543	CEC = −25.82 + 1.01*Clay − 0.037*Sand + 0.543*pH + 0.405*OC + 0.0023*Clay^2 + 0.0057*Sand^2 + 0.0058*Silt^2 + 0.0007*Sand^2 − 0.158*pH^2 − 0.015*OC^2 + 4.337*exp(OC/Sand)^2
ANNs	The entire date set	170 (50 for testing)	0.82	0.84	5.53	

soils have been ploughed for millennia, leading to soil erosion and degradation (Ibáñez et al., 2015).

The results for the pasture categories and for the USA and Iraq were consistent with previous studies. This indicates that CEC in these categories would be largely controlled by parent materials (Nadeu et al., 2012; Mousavi, 2012) through their control of soil texture and mineralogy. However, the role of human activities, such as land use changes, is extremely important in driving soil degradation and can lead to deterioration of soil CEC over both the long and short term (Jalali, 2013; Khaledian et al., 2016a).

Although there are a large number of studies indicating that CEC is controlled by OC content (McBratney et al., 2002; Ulusoy et al., 2016; Khaledian et al., 2016a), the results of the entire data set in this study did not support that. In general, CEC in this study was positively correlated to clay and negatively to sand and pH. It is possible that the depth of sampling played a role in this. Many studies that investigate CEC only sample to about 0–20 cm depth (Laird et al., 2010; Halvorson et al., 2011; Kweon et al., 2013). Soil OC tends to be concentrated in these shallow layers and decrease rapidly with depth but also varies considerably from place to place even in surface soils (Parras-Alcántara et al., 2015). This combination of relatively high concentrations in the surface soil layers, the naturally high CEC of soil organic matter, and a high natural variability in surface soils gives OC a great opportunity to significantly contribute to variations in the CEC of surface soils. However, at some of the sites in this study, such as in Spain, soil OC was low throughout the soil profile, and soils were sampled up to 1 m, depths at which factors other than OC will have a strong influence on CEC. This combination gave other factors that influence soil CEC an opportunity to take on additional importance in the data set.

The accuracy indicators in this study are as promising or more promising than other studies that have predicted CEC using various statistical approaches, such as general linear models (R² ranged between 0.43 and 0.78) (Seybold et al., 2005), multiple linear regression (R² = 0.57 and RMSE = 4.6) (Shabani and Norouzi, 2015), ANNs and ANFIS (R² = 0.88 and RMSE = 1.01) (Ghorbani et al., 2015), PLS (R² = 0.80) and stepwise regression (R² = 0.49) (Khaledian et al., 2016a), genetic expression programming (GEP) (R² = 0.80) and multivariate adaptive regression splines (R² = 0.86) (MARS) (Emamgolizadeh et al., 2015). This indicates that the soil modeling applications of the statistical techniques tested in this study warrant further investigation.

5. Conclusion

Since CEC is an important indicator of soil quality that can be influenced by land use and management and is a soil property which is relatively expensive and time consuming to determine, it is a prime soil property to be estimated using pedotransfer functions. This study was the first to attempt to estimate soil CEC using the PSO and MONANOVA statistical methods and compare these with previously tested models, such as ANNs and PCR. MONANOVA, ANNs, and PSO were able to accurately estimate CEC using clay, sand, pH and OC. Even though the results of MONANOVA and ANNs were more powerful than those from PSO and PCR, PSO provides a formula, which would be necessary in order to use these results to predict CEC in other locations. The relative contribution of the soil properties studied was different in different land uses and countries. We concluded that in each land use and in Iran and Iraq, clay and sand were the most important soil properties determining soil CEC, with clay positively and sand negatively influencing CEC. In Spain OC had the greatest influence on CEC, and clay, silt, sand and OC all had significant influence in the USA. Consequently, this study demonstrated that CEC can be influenced by different soil properties, management and geographical locations. Developing a universally applicable global formula is not likely or will involve a complex formula that will be able to take all these variables into account in an economical fashion.

References

Amini, M., Abbaspour, K.C., Khademi, H., Fathianpour, N., Afyuni, M., Schulin, R., 2005. Neural network models to predict cation exchange capacity in arid regions of Iran. Eur. J. Soil Sci. 53, 748–757.

Behera, S.K., Shukla, A.K., 2015. Spatial distribution of surface soil acidity, electrical conductivity, soil organic carbon content and exchangeable potassium, calcium and magnesium in some cropped acid soils of India. Land Degrad. Dev. 26 (1), 71–79. http://dx.doi.org/10.1002/ldr.2306.

Borggaard, O.K., Szilas, C., Gimsing, A.L., Rasmussen, L.H., 2004. Estimation of soil phosphate adsorption capacity by means of a pedotransfer function. Geoderma 118, 55–61.

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

Brevik, E.C., 2009. Soil health and productivity. In: Verheye, W. (Ed.), Soils, Plant Growth and Crop Production. Encyclopedia of Life Support Systems (EOLSS). EOLSS Publishers, Oxford, UK Developed under the Auspices of the UNESCO. http://www.eolss.net (Accessed 5 Dec., 2015).

Brevik, E.C., Sauer, T.J., 2015. The past, present, and future of soils and human health studies. Soil 1, 35–46. http://dx.doi.org/10.5194/soil-1-35-2015.

Busenberg, E., Clemency, C.V., 1973. Determination of the cation exchange capacity of clays and soils using an ammonia electrode. Clay Clay Miner. 21, 213–217.

Chapman, H.D., 1965. Cation exchange capacity. In: Black, C.A. (Ed.), Methods of Soil Analysis, Part 2, 2nd (ed.). ASA - SSSA No. 9pp. 891–901 Madison, WI.

Chen, H., Lu, X., Li, L.Y., Gaoa, T., Chan, Y., 2014. Metal contamination in campus dust of Xi'an, China: a study based on multivariate statistics and spatial distribution. Sci. Total Environ. 484, 27–35.

Cornelis, W.M., Ronsyn, J., Van Meirvenne, M., Hartmann, R., 2001. Evaluation of pedotransfer functions for predicting the soil moisture retention curve. Soil Sci. Soc. Am. J. 65 (3), 638–648.

Costa, J.L., Aparicio, V., Cerdà, A., 2015. Soil physical quality changes under different management systems after 10 years in the Argentine humid pampa. Solid Earth 6 (1), 361–371. http://dx.doi.org/10.5194/se-6-361-2015.

Emamgolizadeh, S., Bateni, S.M., Shahsavani, D., Ashrafi, T., Ghorbani, H., 2015. Estimation of soil cation exchange capacity using genetic expression programming (GEP) and multivariate adaptive regression splines (MARS). J. Hydrol. 529 (3), 1590–1600. http://dx.doi.org/10.1016/j.jhydrol.2015.08.025.

Gee, G.W., Bauder, J.W., 1986. Particle-size analysis. In: Klute, A. (Ed.), Methods of Soil Analysis. Part I: Physical and Mineralogical Methods, 2nd (ed.). ASA - SSSA No. 9pp. 383–411 Madison, WI.

Ghorbani, H., Kashi, H., Hafezi Moghadas, N., 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. Commun. Soil Sci. Plant Anal. 46, 763–780. http://dx.doi.org/10.1080/00103624.2015.1006367.

Halvorson, J.J., Gonzalez, J.M., Hagerman, A.E., 2011. Repeated applications of tannins and related phenolic compounds are retained by soil and affect cation exchange capacity. Soil Biol. Biochem. 43, 1139–1147. http://dx.doi.org/10.1016/j.soilbio.2011.01.023.

Hosseini, M., Agereh, S.R., Khaledian, Y., Zoghalchali, H.J., Brevik, E.C., Naeini, S.A.R.M., 2017. Comparison of multiple statistical techniques to predict soil phosphorus. App. Soil Ecol. 114, 123–131. http://dx.doi.org/10.1016/j.apsoil.2017.02.011.

Ibáñez, J.J., Pérez-Gómez, R., Oyonarte, C., Brevik, E.C., 2015. Are there arid land soilscapes in southwestern Europe? Land Degrad. Dev. 26 (8), 785–862. http://dx.doi.org/10.1002/ldr.2451.

Jalali, M., 2013. Using chemical analysis and modeling to enhance the understanding of soil solution of some calcareous soils. Environ. Earth Sci. 68 (7), 2041–2049.

Jolliffe, I.T., 1982. A note on the use of principal components in regression. J. R. Stat. Soc. Ser. C 31 (3), 300–303. http://dx.doi.org/10.2307/2348005.

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., Fresco, L.O., 2016. The significance of soils and soil science towards realization of the United Nations sustainable development goals. Soil 2, 111–128. http://dx.doi.org/10.5194/soil-2-111-2016.

Kennedy, J., Eberhart, R.C., 1995. Particle swarm optimization. In: Proceedings of IEEE International Conference on Neural Networks. IEEE Press, Piscataway, NJ, pp. 1942–1948.

Khaledian, Y., Kiani, F., Ebrahimi, E., 2012. The effect of land use change on soil and water quality in northern Iran. J. Mt. Sci. 9 (6), 798–816. http://dx.doi.org/10.1007/s11629-012-2301-1.

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

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

Kruskal, J.B., 1965. Analysis of factorial experiments by estimating monotone transformations of the data. J. R. Stat. Soc. Ser. B 27, 251–263.

Kweon, G., Lund, E., Maxton, C., 2013. Soil organic matter and cation-exchange capacity sensing with on-the-go electrical conductivity and optical sensors. Geoderma 199, 80–89.

Laird, D., Fleming, P., Wang, B., Horton, R., Karlen, D., 2010. Biochar impact on nutrient leaching from a Midwestern agricultural soil. Geoderma 158, 436–442. http://dx.doi.org/10.1016/j.geoderma.2010.05.012.

Mahmoud, E.K., Ghoneim, A.M., 2016. Effect of polluted water on soil and plant contamination by heavy metals in El-Mahla El-Kobra, Egypt. Solid Earth 7, 703–711. http://dx.doi.org/10.5194/se-7-703-2016.

Martínez-Casasnovas, J.A., Ramos, M.C., Benites, G., 2016. Soil and water assessment tool soil loss simulation at the sub-basin scale in the Alt Penedès-Anoia vineyard region (NE Spain) in the 2000s. Land Degrad. Dev. 27 (2), 160–170. http://dx.doi.org/10.1002/ldr.2240.

Masto, R.E., Sheik, S., Nehru, G., Selvi, V.A., George, J., Ram, L.C., 2015. Assessment of environmental soil quality around Sonepur Bazari mine of Raniganj coalfield, India. Solid Earth 6 (3), 811–821. http://dx.doi.org/10.5194/se-6-811-2015.

McBratney, A.B., Odeh, I.O., 1997. Application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions. Geoderma 77, 85–113.

McBratney, A.B., Minasny, B., Cattle, S.R., Vervoort, R.W., 2002. From pedotransfer function to soil inference system. Geoderma 109, 41–73.

McCulloch, W., Pitts, W., 1943. A logical calculus of ideas immanent in nervous activity. Bull. Math. Biophys. 5 (4), 115–133. http://dx.doi.org/10.1007/BF02478259.

McDonald, K.B., 1998. Development of pedotransfer functions of southern Ontario soils. In: Report from Greenhouse and Processing Crops Research Center. Harrow, Ontario, pp. 1–23 No.: 01686-8-0436.

McLean, E.O., 1982. Soil pH and lime requirement. In: Page, A.L., Mille, R., Keeney, D.R. (Eds.), Methods of Soil Analysis, Part 2. Chemical and Microbiological Properties, 2nd Ed. ASA - SSSA No. 9pp. 199–224 Madison, WI.

Mol, G., Keesstra, S., 2012. Soil science in a changing world. Curr. Opin. Environ. Sustain. 4 (5), 473–477. http://dx.doi.org/10.1016/j.cosust.2012.10.013.

Mousavi, S.M., 2012. Application of Rosgen classification the Ziarat watershed, Gorgan. J. Am. Sci. 8 (4), 184–189.

Mukherjee, A., 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.

Nadeu, E., Berhe, A.A., Vente, J.D., Boix-Fayos, C., 2012. Erosion, deposition and replacement of soil organic carbon in Mediterranean catchments: a geomorphological, isotopic and land use change approach. Biogeosciences 9, 1099–1111. http://dx.doi.org/10.5194/bg-9-1099-2012.

Nelson, D.W., Sommers, L.E., 1982. Total carbon, organic carbon, and organic matter. In: Page, A.L., Miller, R.H., Keeney, D.R. (Eds.), Methods of Soil Analysis, Part 2. Chemical and Microbiological Properties, 2nd (ed.). ASA - SSSA No. 9pp. 539–579 Madison, WI.

Page, A.L., Miller, R.H., Keeney, D.R., 1982. Methods of soil analysis. Part 2. In: Chemical and Microbiological Properties, 2nd (ed.). ASA, SSSA, CSSA No. 9pp. 595–623 Madison, WI.

Parras-Alcántara, L., Lozano-García, B., Brevik, E.C., Cerdà, A., 2015. Soil organic carbon stocks assessment in Mediterranean natural areas: a comparison of entire soil profiles and soil control sections. J. Environ. Manag. 155, 219–228. http://dx.doi.org/10.1016/j.jenvman.2015.03.039.

Paz-Ferreiro, J., Fu, S., 2016. Biological indices for soil quality evaluation: perspectives and limitations. Land Degrad. Dev. 27, 14–25.

Pulido, M.M., Gabriels, D., Cornelis, W., Lobo, D., 2015. Comparing aggregate stability tests for soil physical quality indicators. Land Degrad. Dev. 26 (8), 843–852. http://dx.doi.org/10.1002/ldr.2225.

Reidy, B., Simo, I., Sills, P., Creamer, R.E., 2016. Pedotransfer functions for Irish soils – estimation of bulk density (ρ_b) per horizon type. Soil 2, 25–39. http://dx.doi.org/10.5194/soil-2-25-2016.

Rodrigo-Comino, J., Seeger, M., Senciales, J.M., Ruiz-Sinoga, J.D., Ries, J.B., 2016. Spatial and temporal variation of soil hydrological processes on steep slope vineyards (Ruwel-Mosel Valley, Gemany). Cuadernos De Investigacion Geografica 42 (1), 281–306. http://dx.doi.org/10.18172/cig.2934.

Saidi, D., 2012. Importance and role of cation exchange capacity on the physicals properties of the Cheliff saline soils (Algeria). Procedia Eng. 33, 435–449.

Sequeira, C.H., Wills, S.A., Seybold, C.A., West, L.T., 2014. Predicting soil bulk density for incomplete databases. Geoderma 213, 64–73.

Seybold, C.A., Grossman, R.B., Reinsch, T.G., 2005. Predicting cation exchange capacity for soil survey using linear models. Soil Sci. Soc. Am. J. 69, 856–863. http://dx.doi.org/10.2136/sssaj2004.0026.

Shabani, A., Norouzi, M., 2015. Predicting cation exchange capacity by artificial neural network and multiple linear regression using terrain and soil characteristics. Indian J. Sci. Technol. 8 (28). http://dx.doi.org/10.17485/ijst/2015/v8i28/83328.

Smith, P., Cotrufo, M.F., Rumpel, C., Paustian, K., Kuikman, P.J., Elliott, J.A., McDowell, R., Griffiths, R.I., Asakawa, S., Bustamante, M., House, J.I., Sobocká, J., Harper, R., Pan, G., West, P.C., Gerber, J.S., Clark, J.M., Adhya, T., Scholes, R.J., Scholes, M.C., 2015. Biogeochemical cycles and biodiversity as key drivers of ecosystem services provided by soils. Soil 1, 665–685. http://dx.doi.org/10.5194/soil-1-665-2015.

Taghizadeh-Mehrjardi, R., 2016. Digital mapping of cation exchange capacity using genetic programming and soil depth functions in Baneh region, Iran. Arch. Agron. Soil Sci. 62 (1), 109–126.

Taguas, E.V., Guzmán, E., Guzmán, G., Vanwalleghem, T., Gómez, J.A., 2015. Characteristics and importance of rill and gully erosion: a case study in a small catchment of a marginal olive grove. Cuadernos De Investigacion Geografica 41 (1), 107–126. http://dx.doi.org/10.18172/cig.2644.

Ulery, A.L., Graham, R.C., Goforth, B.R., Hubbert, K.R., 2017. Fire effects on cation exchange capacity of California forest and woodland soils. Geoderma 286, 125–130.

Ulusoy, Y., Tekin, Y., Tümsavas, Z., Mouazen, A.M., 2016. Prediction of soil cation exchange capacity using visible and near infrared spectroscopy. Biosyst. Eng. 152, 79–93. http://dx.doi.org/10.1016/j.biosystemseng.2016.03.005.

Van Hall, R.L., Cammeraat, L.H., Keesstra, S.D., Zorn, M., 2017. Impact of secondary vegetation succession on soil quality in a humid Mediterranean landscape. Catena 149, 836–843. http://dx.doi.org/10.1016/j.catena.2016.05.021.

Van Leeuwen, J.P., Lehtinen, T., Lair, G.J., Bloem, J., Hemerik, L., Ragnarsdóttir, K.V.,

Gísladóttir, G., Newton, J.S., de Ruiter, P.C., 2015. An ecosystem approach to assess soil quality in organically and conventionally managed farms in Iceland and Austria. Soil 1, 83–101. http://dx.doi.org/10.5194/soil-1-83-2015.

Willaarts, B.A., Oyonarte, C., Muñoz-Rojas, M., Ibáñez, J.J., Aguilera, P.A., 2016. Environmental factors controlling soil organic carbon stocks in two contrasting Mediterranean-climate areas. Land Degrad. Dev. 27, 603–611. http://dx.doi.org/10.1002/ldr.2417.

Yukselen, Y., Kaya, A., 2008. Suitability of the methylene blue test for surface area, cation exchange capacity and swell potential determination of clayey soils. Eng. Geol. 102, 38–45.

Zeraatpishe, M., Khormali, F., 2012. Carbon stock and mineral factors controlling soil organic carbon in a climatic gradient, Golestan Province. J. Soil Sci. Plant Nutr. 12 (4), 637–654.

Zolfaghari, A.A., Taghizadeh-Mehrjardi, R., Moshki, A.R., Malone, B.P., Weldeyohannes, A.O., Sarmadian, F., Yazdani, M.R., 2016. Using the nonparametric k-nearest neighbor approach for predicting cation exchange capacity. Geoderma 265, 111–119.