Applied Clay Science 116–117 (2015) 129–140

Contents lists available at ScienceDirect

Applied Clay Science

journal homepage: www.elsevier.com/locate/clay

Research paper

Analyzing the effect of various soil properties on the estimation of soil specific surface area by different methods

Hossein Bayat [a,*], Eisa Ebrahimi [b], Sabit Ersahin [c], Estela N. Hepper [d], Devendra Narain Singh [e], Abdel-monem Mohamed Amer [f], Yeliz Yukselen-Aksoy [g]

[a] Department of Soil Science, Faculty of Agriculture, Bu Ali Sina University, Hamedan, Iran
[b] Young Researches and Elite Club, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran
[c] Department of Soil Science, Agricultural Faculty, Gaziosmanpasa University, 60250 Taslıciftlik Tokat, Turkey
[d] Facultad de Agronomía, UNLPam, cc 300, 6300 Santa Rosa, Argentina
[e] Department of Civil Engineering, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India
[f] Soil Science Department, Faculty of Agriculture, Menoufia University, Shebin El-Kom,Egypt
[g] Department of Civil Engineering, Celal Bayar University, Muradiye Kampusu, Muradiye Manisa, Turkey

ARTICLE INFO

Article history:
Received 6 November 2014
Received in revised form 16 June 2015
Accepted 24 July 2015
Available online 28 August 2015

Keywords:
Artificial neural network
Group method of data handling
Pedotransfer functions
Regression trees
Specific surface area
Sensitivity analysis

ABSTRACT

Depending on the method used, measuring the specific surface area (SSA) can be expensive and time consuming and limited numbers of studies have been conducted to predict SSA from soil properties. In this study, 127 soil sample data were gathered from the available literature. The data set included SSA values and some of the soil physical and chemical index properties. At the first step, linear regression, non-linear regression, regression trees, artificial neural networks, and a multi-objective group method of data handling were used to develop seven pedotransfer functions (PTFs) for the purpose of finding the best method in predicting SSA. Results showed that the artificial neural networks performed better than the other methods used in the development and validation of PTFs. At the second step, to find the best set of SSA for predicting input variables and to investigate the importance of the input parameters, the artificial neural networks were further used and 25 models were developed. The results showed that the PTF, containing the input variables of sand%, clay%, plastic limit, liquid limit, and free swelling index performed better than the other PTFs. This can be attributed to the close relation between the free swelling index and Atterberg limits with the soil clay mineralogy, which is one of the most important factors controlling SSA. The sensitivity analysis showed that the greatest sensitivity coefficients were found for the cation exchange capacity, clay content, liquid limit, and plasticity index in different models. Overall, the artificial neural networks method was proper to predict SSA from soil variables.

1. Introduction

Soil specific surface area (SSA) is a fundamental soil property that can be used as an indicator of soil behavior (Utkaeva, 2007) to explain many of the physical and chemical phenomena of the soil: including fertility-determining components (such as the water holding capacity, the adsorption of plant nutrients, and the amount of organic matter) (Voronin, 1975; Shein, 2005). The SSA is strongly related to the dry bonding power of clayey soils (Yukselen and Kaya, 2008), the adsorption of polar compounds such as pesticides and pollutants, swell-shrinkage behavior (Yukselen and Kaya, 2008), formation of the soil structure and quality of the resulting soil aggregates (Utkaeva, 2007), ion exchange properties of clay minerals (Salehi et al., 2008), contaminant transport, hydraulic conductivity (Altin et al., 1999), biological processes (Rawlins et al., 2010), cation exchange capacity (Petersen et al., 1996), Atterberg limits of fine-grained soils (Dolinar et al., 2007), water retention (Tuller and Or, 2005), release rate of base cations (Hodson et al., 1998) and groundwater vulnerability (Maxe and Johansson, 1998). In addition, there are strong correlations between SSA and soil texture, $CaCO_3$ content, and the salinity and mineralogical composition of the soil, which refers to the absence or presence of internal pores (Monem and Amer, 2009).

The SSA highly depends on the measurement method due to the materials used in each method; and the results obtained are questionable

Abbreviations: S, Sand content; Si, Silt content; Cl, Clay content; SSA, Specific Surface Area; OM, Organic Matter; CEC, Cation Exchange Capacity; LL, Liquid Limit; PL, Plastic Limit; SL, Shrinkage Limit; PI, Plasticity Index; FSI, Free Swelling Index; n, number of samples used; R, Regression; NL, Non-linear Regression; ANN, artificial neural network; GMDH, Group method of data handling; RT, Regression Trees; AIC, Akaike Information Criterion; RMSE, Root Mean Square Error; MGN, Morgan–Granger–Newbold.

* Corresponding author.
E-mail addresses: h.bayat@basu.ac.ir, hbayat2001@gmail.com (H. Bayat), ebrahimi.soil@gmail.com (E. Ebrahimi), acapsu@gmail.com (S. Ersahin), hepper@agro.unlpam.edu.ar (E.N. Hepper), dns@civil.iitb.ac.in (D.N. Singh), amerabdel@hotmail.com (A.M. Amer), yeliz.aksoy@bayar.edu.tr (Y. Yukselen-Aksoy).

http://dx.doi.org/10.1016/j.clay.2015.07.035
0169-1317/

and inconsistent (Chiou and Rutherford, 1993; Churchman et al., 1991; de Jong, 1999; Hang and Brindley, 1970; Yukselen and Kaya, 2006). The reported SSA of expanded clays ranged between 560 and 800 m^2/g measured with ethylene glycol mono ethyl ether (Heilman et al., 1965); and the external SSA of clay minerals ranged from 10 to 20 m^2/g measured by N_2-adsorption technique (Altin et al., 1999). In addition, measurements of SSA are relatively expensive, time consuming, and labor intensive (Rawlins et al., 2010). Therefore, prediction methods are needed for an effective determination of the soil SSA, from routinely measured soil variables (Moiseev, 2008).

There are limited attempts that have been made to develop pedotransfer functions (PTFs) to describe SSA from readily available soil variables. Most of these PTFs were regression equations developed to predict SSA from related soil properties (Ustinov and Kuznetsov, 1982; Sapozhnikov et al., 1992). Recently, Bayat et al. (2013) have used artificial neural networks (ANNs) to estimate SSA. Also, Ismeik and Al-Rawi (2014) used some of geotechnical properties of soils as predictors to estimate SSA by ANNs and reported reliable results. According to Salehi et al. (2008), PTFs predict difficult to measure soil properties, such as SSA, from more readily available soil properties (e.g., hygroscopic water content, bulk density, etc.). It has also been suggested that a simple empirical equation can describe the relationship between mineral surface area and soil particle-size distribution (Sverdrup et al., 1990). In addition to empirical regression-based model such as PTFs, Tuller and Or (2005) proposed a method based on van der Waals adsorbed water films in which SSA is used to predict the very dry end of soil water retention curve. This approach can be inversely applied to estimate SSA from the dry end of soil water retention curve.

Different variables have been used to estimate SSA, such as V (Vanadium), Ca (Calcium), Al (Aluminum), and Rb (Rubidium) (Rawlins et al., 2010), the maximum hygroscopic water content (using Mitscherlich equation) (Moiseev, 2008), the soil water retention data corresponding to matric potentials of less than −10 MPa (Resurreccion et al., 2011), soil water content (Moiseev, 2008), soil texture, organic matter, and fractal parameters of particle size distribution (Bayat et al., 2013).

ANNs have performed better than the other techniques and they have overcome the problem of introducing statistical assumptions into PTFs (Pachepsky et al., 1996; Schaap and Bouten, 1996; Tamari et al., 1996; Koekkoek and Booltink, 1999; Minasny et al., 1999; Minasny and McBratney, 2002). Moreover, other computer techniques such as multi-objective group method of data handling (mGMDH) (Bayat et al., 2011) and regression trees (RT) (Rawls et al., 2003; Toth et al., 2012) have been used to develop PTFs to estimate soil hydraulic properties.

Although many efforts have been made to develop PTFs using regression methods, the study of ANNs to estimate SSA conducted by Bayat et al. (2013) and Ismeik and Al-Rawi (2014) were the only works of its kind, and to our knowledge no research has been conducted to evaluate the performance of mGMDH and RT in estimating SSA from readily available soil data.

Despite the diverse parameters that have been used in developing PTFs, two basic questions still remain unanswered about the estimation of SSA; the first one is about which input variables are preferable or necessary to estimate SSA and whether new variables can be found to improve the performance of PTFs or not; and the second one is which methods are the most appropriate to develop a PTF to estimate SSA.

The aims of this study were (i) to develop PTFs to estimate SSA using various techniques, such as simple linear regression (R), non-linear regression (NL), ANNs, mGMDH, and RT; and (ii), to evaluate the utility of different predictors in the prediction of SSA and to find the best subsets of predictors.

2. Materials and methods

2.1. Data sets

In total, 127 soil samples collected from 8 available databases in the literature as follows: Aringhieri et al. (1992) (17 samples), Ersahin et al. (2006) (21 samples), Hepper et al. (2006) (24 samples), Yukselen and Kaya (2006) (1 sample), Arnepalli et al. (2008) (4 samples), Amer (2009) (22 samples), Yukselen and Kaya (2010) (12 samples), and Srinivas (submitted for publication) (26 samples). Table 1 shows the detailed information about the datasets used in this study. The textures of the studied soils are shown in Fig. 1. Most of the soils took place in the clay textural class and no soil sample fell in sandy clay and silt textures.

Sand content S(%), silt content Si(%), clay content Cl(%) (ASTM D-4221994; Schlichting et al., 1995; Gee and Bouder, 1986), organic matter, OM(%) (Walkley and Black, 1934; ASTM, 1999), cation exchange capacity CEC ($cmol_c$/kg) (IS, 27201976; Rhoades, 1982), liquid limit LL(%), plastic limit PL(%), shrinkage limit SL(%), plasticity index PI(%) (ASTM, 1999), free swelling index FSI(%) (Thakur and Singh, 2005; Shah, submitted for publication) and SSA were the soil properties selected to develop PTFs in this study. Aringhieri et al. (1992) studied soils that had a wide range of clay mineralogy, iron oxides, amorphous alumina-silicate and OM. The soil samples that have been studied by Ersahin et al. (2006) were taken from distinct parent materials under diverse climate, vegetation, and topography. Hepper et al. (2006) studied soils from 12 sites of the semiarid Argentinean Pampean Area including ash free (sites 1 to 5) and ash enriched (sites 6 to 12) soils. Yukselen and Kaya (2006, 2010) studied soils with different origins and characteristics: all samples were obtained from different parts of Turkey, except one.

Silty-soil, clayey silt, a vertisol, and a kaolinitic soil studied by Arnepalli et al. (2008), were selected for this study. Four soils, which were different in their texture, salinity, and $CaCO_3$ were chosen from the dataset of Amer (2009). Two of these four soils are typical calcareous soils from Borg El-Arab (>30% $CaCO_3$) and El-Nubaria (<30% $CaCO_3$) areas (Amer, 2009). Texture of these soils are loamy sand and sandy clay loam, respectively. The other two soils are a non-saline alluvial clay soil and a saline alluvial clay soil, both had been taken from Epshan and El-Khamsen, Kafr El-Sheikh Province. Details of these four soils and their laboratory analyses can be found in Amer (2009).

Ersahin et al. (2006) measured the SSA of the soils, analyzing the retention level of ethylene glycol mono ethylene ether (EGME), which is a polar molecule that forms only one layer of molecules on the particle surfaces (Cerato and Lutenegger, 2002). Hepper et al. (2006) determined the SSA with EGHE after sieving the samples (<0.5 mm), peroxidation, saturation with Ca^{2+}, and air drying (Carter et al., 1986). Aringhieri et al. (1992) used the method of Quirk (1955) to measure SSA by water vapor adsorption. Amer (2009) measured the total SSA applying the Brunauer, Emmett and Teller (BET)- N_2 adsorption method (Brunauer et al., 1938) as modified and described by Orchiston (1954), Quirk (1955), Farrar (1963), and Globus (1996). Arnepalli

Table 1
Number of data that was collected from each database for every variable.

Sources	S[a]	Si	Cl	SSA	OM	CEC	LL	PL	SL	PI	FSI
Ersahin et al. (2006)	21	21	21	21	21	21	–	–	–	–	–
Hepper et al. (2006)	24	24	24	24	24	24	–	–	–	–	–
Aringhieri et al. (1992)	17	17	17	17	17	17	–	–	–	–	–
Amer (2009)	22	22	22	22	22	–	–	–	–	–	–
Yukselen and Kaya (2010)	12	12	12	12	12	12	12	11	12	11	12
Yukselen and Kaya (2006)	1	1	1	1	1	1	1	1	1	1	1
Srinivas (submitted for publication)	26	26	26	26	3	3	25	25	24	25	24
Arnepalli et al. (2008)	4	4	4	4	–	4	4	4	4	4	4

[a] Sand content, S, silt content, Si, clay content, Cl, specific surface area, SSA, organic matter, OM, cation exchange capacity, CEC, liquid limit, LL, plastic limit, PL, shrinkage limit, SL, plasticity index, PI, free swelling index, FSI.