Catena 111 (2013) 72–79

Contents lists available at ScienceDirect

Catena

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

Estimating wet soil aggregate stability from easily available properties in a highly mountainous watershed

A.A. Besalatpour [a,*], S. Ayoubi [b], M.A. Hajabbasi [b], M.R. Mosaddeghi [b], R. Schulin [c]

[a] Department of Soil Science, College of Agriculture, Vali-e-Asr University of Rafsanjan, 7718897111 Rafsanjan, Iran
[b] Department of Soil Science, College of Agriculture, Isfahan University of Technology, 84156-83111 Isfahan, Iran
[c] Institute of Terrestrial Ecosystems, ETH Zurich, Universitätstr 16, CH-8092 Zürich, Switzerland

ARTICLE INFO

Article history:
Received 25 March 2013
Received in revised form 25 June 2013
Accepted 1 July 2013

Keywords:
Intelligence systems
Neuro-fuzzy inference system
Artificial neural networks
Conventional regression models
Mean weight diameter

ABSTRACT

A comparison study was carried out with the purpose of verifying when the adaptive neuro-fuzzy inference system (ANFIS), artificial neural network (ANN), generalized linear model (GLM), and multiple linear regression (MLR) models are appropriate for prediction of soil wet aggregate stability (as quantified by the mean weight diameter, MWD) in a highly mountainous watershed (Bazoft watershed, southwestern Iran). Three different sets of easily available properties were used as inputs. The first set (denoted as SP) consisted of soil properties including clay content, calcium carbonate equivalent, and soil organic matter content. The second set (denoted as TVA) included topographic attributes (slope and aspect) and the normalized difference vegetation index (NDVI). The third set (denoted as STV) was a combination of soil properties, slope, and NDVI. The ANN and ANFIS models predicted MWD more accurately than the GLM and MLR models. Estimation of MWD using TVA data set resulted in the lowest model efficiency values. The observed model efficiency values for the developed MLR, GLM, ANN, and ANFIS models using the SP data set were 60.76, 62.98, 77.68 and 77.15, respectively. Adding slope and NDVI to soil data (i.e. STV data set) improved the predictions of all four methods. The obtained correlation coefficient values between the predicted and measured MWD for the developed MLR, GLM, ANN, and ANFIS models using STV data set were 0.24, 0.35, 0.84 and 0.73, respectively. In conclusion, the ANN and ANFIS models showed greater potential in predicting soil aggregate stability from soil and site characteristics, whereas linear regression methods did not perform well.

1. Introduction

Soil aggregate stability is a key factor of soil resistivity to mechanical stresses, including the impacts of rainfall and surface runoff, and thus to water erosion (Canasveras et al., 2010). When soil aggregates break down, finer particles are produced, which are easily carried away by wind and water flow and which upon re-sedimentation tend to clog soil pores, leading to the formation of soil crusts (Kirkby and Morgan, 1980; Renard et al., 1997; Yan et al., 2008). Reducing infiltration, this sealing effect enhances surface runoff and thus promotes further water erosion. Hence, aggregate stability is an important factor in soil erosion.

Various indicators have been proposed to characterize and quantify soil aggregate stability, for example percentage of water-stable aggregates (WSA), mean weight diameter (MWD) and geometric mean diameter (GMD) of aggregates, and water-dispersible clay (WDC) content (Calero et al., 2008; Le Bissonnais, 1996). Unfortunately, the experimental methods available to determine these indicators are laborious, time-consuming and difficult to standardize (Canasveras et al., 2010). Therefore, it would be advantageous if aggregate stability could be predicted indirectly from more easily available data.

General soil properties most closely correlated with soil aggregate stability are the contents of clay, calcium carbonate, and organic matter (Canasveras et al., 2010). Clay particles are considered as cementing agents for aggregation because of their high specific surface area, high cation exchange capacity (CEC), and consequently, high physical and chemical activity. Soil organic matter content can affect soil structure as well as soil aggregate stability in different ways: the transient aggregating effect of polysaccharides on micro-aggregates, increased aggregate coherence against slaking due to hydrophobic materials, the temporarily stabilizing effect of roots and hyphae on macro-aggregates, and the persistent effect of polymers and aromatic compounds on micro-aggregates. Calcium carbonate contents also influence soil aggregation through their cementing effects and preventing aggregate dispersion (Amezketa, 1999).

Indirectly, also topography and vegetation characteristics affect aggregate stability, in particular through their influence on the dynamics of soil structure and soil properties such as clay mineralogy,

* Corresponding author. Tel.: +98 913 1670128; fax: +98 391 3202042.
E-mail addresses: a.besalatpour@vru.ac.ir, a_besalatpour@yahoo.com (A.A. Besalatpour).

0341-8162 © 2013 The Authors. Published by Elsevier B.V. Open access under CC BY-NC-ND license.
http://dx.doi.org/10.1016/j.catena.2013.07.001

SOM, carbonate concentration, texture, soil water content, and plant development (Canton et al., 2009). Furthermore, slope and aspect may influence the rate of weathering and erodibility of soils and thus soil aggregate stability (Bronick and Lal, 2005).

Functions translating such data into predictions of soil aggregate stability can be derived by a variety of methods. In contrast to widespread applications of conventional regression models to predict soil aggregate stability indirectly from other data; artificial intelligence systems such as artificial neural networks (ANNs) and adaptive neuro-fuzzy inference system (ANFIS) have not been exploited for this purpose, although they have shown much potential in similar applications (Azamathulla et al., 2009; Bocco et al., 2010; Gago et al., 2010; Huading et al., 2007; Huang et al., 2010; Kisi et al., 2009; Uno et al., 2005).

ANNs are computing systems made up of a number of simple, highly interconnected processing elements, also called neurons. Generally, an ANN is made of an input layer, one or several hidden layers (HLs), and an output layer of neurons (Tracey et al., 2011). The input layer neurons receive the input information from the outside environment and transmit it to hidden layer. Each neuron of a subsequent layer first computes a linear combination of the outputs from all neurons of the previous layer and then adds a bias to it. Furthermore, each neuron of a HL applies a specific nonlinear function, called *activation function*, to this linear combination plus bias. The coefficients of the linear combinations and the biases are called weights (Bocco et al., 2010; Saridemir, 2009; Sobhani et al., 2010; Turan et al., 2011).

ANFIS is a scheme that uses the learning capability of ANNs to derive fuzzy IF–THEN rules with appropriate fuzzy set membership functions (Jang and Sun, 1995; Tay and Zhang, 1999). The main strength of ANFIS in comparison with ANNs is that it generates linguistically interpretable IF–THEN rules (Sobhani et al., 2010). ANFIS models capture the relationship between input and output data by establishing fuzzy language rules, while ANNs do so in form of trained connection weights. Furthermore, it is reported that constructing an ANFIS model is less time-consuming than an ANN model (Azamathulla et al., 2009).

The objectives of this study were to: i) compare the capabilities of ANFIS, ANN, generalized linear model (GLM), and multiple linear regression (MLR) to derive pedotransfer functions (PTFs) between soil aggregate stability and various sets of input variables, and ii) use the PTFs for prediction of aggregate stability using another set of soil samples collected from the same area. For this purpose, three different sets of easily available data including soil properties alone, topographic attributes and vegetation index, and a combination of soil properties and topographic and vegetation attributes were used as inputs. The current comparison study in using different soft computing techniques and also different data sets for MWD estimation can be a valuable source of information for other modelers. Discussions of advantages and disadvantages are also given in different point of view for all the methods.

2. Materials and methods

2.1. Study area description

The study area was part of the Bazoft watershed (31° 37′ to 32° 39′ N and 49° 34′ to 50° 32′ E), which is located in the northern part of the Karun river basin in central Iran. The major river in the watershed is the Ab-Bazoft, which joins the Karun River at the outlet of the watershed. The elevation ranges from 880 m a.s.l. in the south of the watershed to 4300 m a.s.l. on the Zardkuh Mountain in the north. The long-term average rainfall of the region varies between 500 and 1400 mm per year, and the average temperature varies between 8 and 20 °C. The watershed is highly mountainous where the most slopes are between 40 and 70%, covering about 46% of the watershed. The dominant slope shape is convex. Approximately 56% of the watershed area is covered by pastures, the rest by forests and bare lands. *Quercus brantii* is the dominating forest tree species, and *Astragalus* sp. is the most abundant pasture plant.

Old terrace deposits (Qt1) are dominant geological unit having moderate susceptibility to weathering and erosion with some marls enrichment with gypsiferous and sandstone (mp1) (Iranian Geological Organization, 2006). Majority of the soils include Calcic Argixerolls, Typic Calcixerepts, Typic Xerorthents, Typic Cryorthents, and Typic Haploxerolls in the watershed (Soil Survey Staff, 2006) which are dominated by calcareous materials. Soils are less than 5 cm deep on steep slopes and more than 150 cm deep in the valley bottoms. The main textural classes are silt loam, loam, silty clay loam, clay loam, and silty clay. The dominant physiographic units are mountains, hills, plateaus and upper traces, alluvial plains, and gravelly colluvial fans.

2.2. Soil sampling and measurements

A stratified random sampling was designed using digital geology, topography, and land use maps in the environment of ILWIS 3.4 software (ITC, University of Twente, Netherlands) for proper selection of soil sampling locations in all of the land uses. Thus, land use type was indirectly taken into account in the soil sampling. In other words, the land use directly or indirectly affects the soil properties (like texture, calcium carbonate, and organic matter) as well as vegetation cover (as quantified by NDVI) which were used as predictors in the PTFs. A total of 160 soil samples were collected from the top 5 cm of soil surface from all major land unit tracts. The positions of the sampling points were identified in the field using GPS (model: 76CSx).

The soil samples were air-dried and ground to pass a 2-mm sieve. Soil organic matter (SOM) content was determined by the Walkley–Black method with dichromate extraction and titrimetric quantization (Nelson and Sommers, 1986). Clay content (<2 μm) was measured by means of sieving and sedimentation using the procedure described by Gee and Bauder (1986), and calcium carbonate equivalent (CCE) was determined by the back-titration method (Nelson, 1982).

The soil samples for aggregate stability assessment were taken from the same locations and brought to the laboratory in such a way that minimum structural deformation and/or destruction occurred. Following van Bavel (1950) method, as modified by Kemper and Rosenau (1986), was used to parameterize the mean weight diameter (MWD) of wet-sieved aggregates. Briefly, 50 g of the <4.75 mm aggregates were placed on the topmost of a stack of sieves with descending mesh size (2, 1, 0.5, and 0.25 mm) from top to bottom. The samples were first immersed in distilled water and then sieved by moving the sieve set vertically. The soil retained by each sieve was dried at 105 °C for 24 h, weighed and corrected for sand/gravel particles to obtain the proportion of water-stable aggregates. The MWD (mm) of water-stable aggregates was calculated using the following equation:

$$\mathrm{MWD} = \sum_{i=1}^{n} w_i \overline{X}_i \tag{1}$$

where $\overline{X}_i$ is the arithmetic mean diameter of each size fraction (mm) and w_i the proportion of the total water-stable aggregates in the corresponding size fraction after deducting the weight of sand/gravel particles (upon dispersion and passing through the same sieve) as indicated above.

2.3. Topographic and vegetation attributes

The topographic attributes, including slope and aspect, were determined using a 20 m by 20 m digital elevation model (DEM). The slope of each cell represents the maximum rate of elevation change between the cell and its neighbor cells. The aspect represents the direction of slope, i.e. of the maximum rate of elevation change in down-slope direction. The normalized difference vegetation index (NDVI) was used to quantify the vegetation cover. It was derived from an IRS-1D satellite photo taken in April 2008 with a spatial resolution of 24 m by 24 m (Indian Space Applications Centre, 2002).