Geoderma 337 (2019) 514–523

Contents lists available at ScienceDirect

Geoderma

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

Performance of the tangential model of soil water retention curves for various soil texture classes

Magatt Thiam[a], Dang Quoc Thuyet[b,c], Hirotaka Saito[a,*], Yuji Kohgo[a]

[a] United Graduate School of Agricultural Sciences, Tokyo University of Agriculture and Technology, 3-5-8, Saiwaicho, Fuchu, Tokyo 183-8509, Japan
[b] Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1, Yayoi, Bunkyo, Tokyo 113-8657, Japan
[c] Institute of Agricultural Machinery, National Agriculture and Food Research Organization, 1-40-2 Nisshin, Kita-ku, Saitama City, Saitama 331-8537, Japan

ARTICLE INFO

Handling Editor: Morgan Cristine L.S.

Keywords:
Soil water retention curve
Soil texture class
Tangential model
Unsaturated soil
Machine learning

ABSTRACT

The tangential model (TANMOD) is one of the few soil water retention curve (SWRC) models that can be applied in both unsaturated and saturated soils, from the positive suction range to the negative suction range, accounting for the effect of volume changes in the entrapped air in soil pores. The model has been successfully evaluated with relatively coarser soils. Its performance, however, has not been fully tested for various soil texture classes. In this study, we aim 1) to determine the TANMOD parameters for various soil texture classes, and 2) to assess the underlying relationship between the TANMOD parameters and the soil texture class. To address those objectives, the TANMOD was first fitted to 399 SWRC from 10 USDA soil texture classes in the UNSODA soil database. The model parameters consist of three coordinates (S_{re}, s_e), (S_{rm}, s_m), and (S_{rf}, s_f), three tangential slopes, c_e, c_m, and c_f, along the curve. Multivariate analysis and several machine learning algorithms were respectively used to evaluate model parameters for each soil texture class and reveal the relation between the model parameters and the soil texture classes. The results demonstrated that the TANMOD fitted well from coarser soils to finer soils. Unique sets of the model parameters and their uncertainties are proposed for 10 USDA soil texture classes. Unsupervised learning algorithms, hierarchical cluster analysis and *k*-means clustering, failed to classify the TANMOD parameters while one of the supervised machine learning techniques, random forest, adequately classified the TANMOD parameters to the USDA soil texture classes. The accuracy of the classification based on the random forest model is 62.6%. The maximum tangential slope, c_m, was the most important parameter in relation with the soil texture class. Consequently, the TANMOD parameters not only have their own physical meaning but also can be applied to various USDA soil texture classes.

1. Introduction

It is important to account not only for negative pore pressure ranges (i.e., positive suction ranges) but also for positive pore pressure ranges (i.e., negative suction ranges) for some physical and mechanical processes of the soil because excess (positive) pore water pressures may play an important role. For example, to analyze dynamic behaviors of soil structures it is necessary to solve field equations with constitutive equations for saturated-unsaturated soils (e.g., Kohgo et al., 2010). The constitutive equations may include non-linear soil water retention, permeability, and stress-strain relationships. Then a soil water retention curve (SWRC) model, which covers the entire potential range from positive to negative, is necessary. There are a number of common SWRC models with a simple functional form with only a few number of parameters, such as the model by Brooks and Corey (1966) or by van Genuchten (1980). Although most parameters used in those models cannot be related directly with physical properties of the soil, in other words, do not have obvious physical meanings, the models have been successfully used in analyzing variably saturated water flow in soils. Those SWRC models are usually combined with a theoretical model of, for example, Mualem (1976) to predict unsaturated hydraulic conductivity. On the other hand, they are not quite suitable when behaviors of unsaturated soil structures are analyzed because they intrinsically assume that the soil is fully saturated in positive pore water pressure (i.e., negative suction) ranges and ignore the effect of entrapped air in soil pores. Kohgo et al. (1993) introduced the concept of "insular" air saturation condition where air exists as bubbles (therefore entrapped) when the suction is smaller than the air entry value. Under this condition, a change in saturation is due to compression of the fluid mixed with air and water in pores. Most available SWRC models cannot

* Corresponding author.
E-mail address: hiros@cc.tuat.ac.jp (H. Saito).

https://doi.org/10.1016/j.geoderma.2018.10.008
Received 29 December 2017; Received in revised form 30 September 2018; Accepted 3 October 2018

account for such conditions. The tangential model (TANMOD) proposed by Kohgo (2008) is one of the few models that can be applied from the negative suction range to the positive suction range accounting for the aforementioned conditions. The model ensures the continuity of the tangential slope of SWRC while accounting for the compression of the air in pores in the insular air saturation condition. The model is also capable of modeling hysteresis (Kohgo, 2008). It has been successfully used in analyzing the behavior of fill-type dams during construction and impounding (Kohgo et al., 2010).

The TANMOD, which has been developed for a soil having a unimodal pore size distribution, requires coordinates of three points and three tangential slopes along the SWRC as inputs (Kohgo, 2008). Then the parameters in this model are numerical values determined from the experimental SWRC. Those three points, referred to as E, M, and F, corresponding to an air entry point, a transitional point with the maximum tangential slope, and a dry end, respectively as shown in Fig. 1. The performance of the TANMOD has not yet been fully tested for various soil texture classes. While the TANMOD parameters are determined based on the shapes of SWRCs, soil texture classes are based on soil particle size distributions. Although the two concepts are different, it is important to investigate if any TANMOD parameter sets can be proposed for different soil texture classes as a number of pedotransfer functions have been proposed for some commonly-used SWRC models, such as by van Genuchten (1980) or by Brooks and Corey (1966) models. It is, therefore, necessary to reveal the relationship between each TANMOD parameter set and the soil texture class. Several approaches have been used to find the connection between SWRC model parameters and soil texture classes. A simple method could be using a lookup table of the soil texture class and the model parameters (Carsel and Parrish, 1988) or averaging the model parameter for each soil class, or using linear and non-linear models (Minasny et al., 1999). Schaap et al. (2001) developed a program called Rosetta to predict van Genuchten model parameters for each soil texture class based on the USDA criteria.

Recently, application of machine learning techniques for classification and modeling has increased. For example, Thuyet et al. (2016) successfully classified trace elements and major irons in groundwater based on aquifer properties using an unsupervised classification, hierarchical cluster analysis (HCA). Twarakavi et al. (2010) were able to use the k-means clustering technique to classify soil hydraulic parameters into so-called soil hydraulic classes, which correspond well with soil texture classes, especially for coarser soils. On the other hand, Cisty et al. (2015) used a random forest algorithm, one of the supervised classification methods, to convert one soil texture classification system to another. The advantage of machine learning techniques as compared to other traditional classification methods is that a given model does not necessarily rely on pre-defined equations.

The main objectives of this study were therefore 1) to determine the TANMOD parameters for various soil texture classes, and 2) to assess the underlying relationship between TANMOD parameters and soil texture classes, as well as the relative importance of each parameter to relate soil texture class with the TANMOD parameters using machine learning techniques. The first objective was achieved by computing mean TANMOD parameters from a set of the TANMOD parameters for each soil texture class. The parameter sets used in this manuscript were obtained by fitting the TANMOD to a number of SWRCs available for 10 soil texture classes in a soil hydraulic database. As for the second objective, the TANMOD parameters were classified into 10 groups using standard classification algorithms. If the TANMOD parameter set in each soil texture class exclusively represents the soil texture class, an appropriate classification algorithm should be able to split the parameters into the corresponding 10 soil texture classes. For each soil texture class, the rates of correct and incorrect classification, as well as, the importance of each model parameter were computed to assess the overall performance of TANMOD for various soil texture classes.

2. Materials and methods

2.1. The tangential model (TANMOD) parameters

The application of the tangential model (TANMOD) to a soil water retention curve (SWRC) requires coordinates of three points, E, M, and F (see Fig. 1) along a given SWRC and three tangential slopes as inputs (Kohgo, 2008). Their coordinates are given as (s_e, S_{re}), (s_m, S_{rm}), and (s_f, S_{rf}), respectively, where the characters, s_i and S_{ri}, respectively, are used for the suction and the saturation. The subscript “i” denotes the point E, M, and F (see Fig. 1). The three so-called tangential slopes, referred to as c_e, c_m, and c_f, correspond to an increase in the saturation with a unit decrease in the suction (i.e., $-\partial S_r/\partial s$) at the insular air saturation condition (beyond E), at the transitional point M within the so-called fuzzy saturation condition, and beyond the dry end point F, referred to as the pendular saturation condition. At the point M, the tangential slope is assumed to be at its maximum (i.e., the change in the saturation with a unit decrease in the suction is the largest). The point F identifies the point at which a large increase in suction would cause an insignificant change in the degree of saturation. The coordinates of the three points, E, M, and F, and the tangential slopes, c_e, c_m, and c_f, are referred to as the TANMOD parameters for convenience in this manuscript. Unlike other SWRC models, there is little risk of misinterpreting the TANMOD parameters. Mathematical expressions to relate the saturation S_r to the suction s; and the slope c_e to S_r are given by Eq. (1) to Eq. (8).

$$S_r = S_{re} + c_e\left(s_e - \frac{P_a s}{P_a - s}\right) \qquad \text{for } s \le 0 \tag{1}$$

$$S_r = S_{re} + c_e(s_e - s) \qquad \text{for } 0 < s \le s_e \tag{2}$$

$$S_r = S_{re} - c_e(s - s_e) - \frac{(c_m - c_e)(s - s_e)^{m_r+1}}{(m_r + 1)(s_m - s_e)^{m_r}} \qquad \text{for } s_e < s < s_m \tag{3}$$

$$S_r = S_{rf} - c_f(s - s_f) - \frac{(c_m - c_f)(s - s_f)^{n_r+1}}{(n_r + 1)(s_m - s_f)^{n_r}} \qquad \text{for } s_m \le s < s_f \tag{4}$$

$$S_r = S_{rf} - c_f(s - s_f) \qquad \text{for } s \ge s_f \tag{5}$$

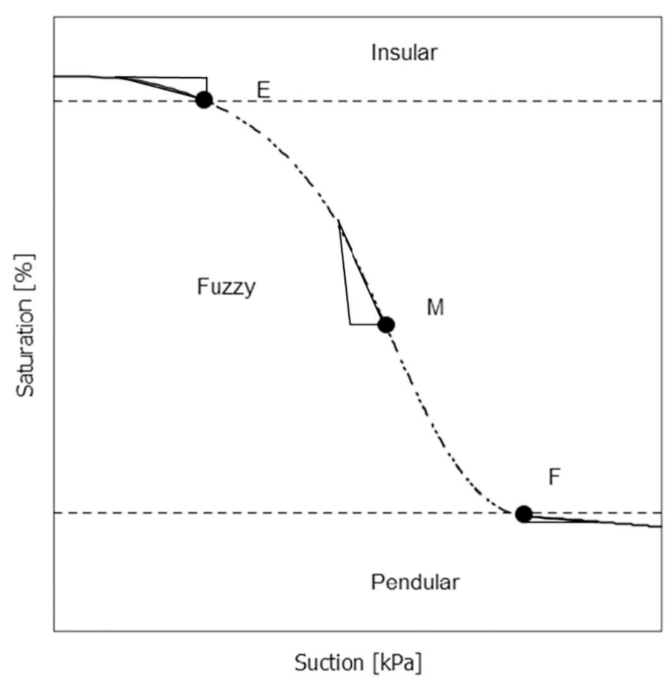

Fig. 1. Schematic representation of the tangential model.