UNIVERSIDADE DE SÃO PAULO

CENTRO DE ENERGIA NUCLEAR NA AGRICULTURA

LEONARDO INFORSATO

An improved methodology for obtaining soil hydraulic properties by laboratory evaporation experiments

Piracicaba

2023

LEONARDO INFORSATO

An improved methodology for obtaining soil hydraulic properties by laboratory evaporation experiments

Thesis presented to the Center for Nuclear Energy in Agriculture at the University of São Paulo to obtain the title of Doctor of Science.

Concentration area: Nuclear Energy in Agriculture and the Environment

Supervisor: Prof. Dr. Quirijn de Jong van Lier

Piracicaba

2023

Dados Internacionais de Catalogação na Publicação (CIP)

Seção Técnica de Biblioteca - CENA/USP (CIP)

Inforsato, Leonardo

Metodologia aperfeiçoada para a obtenção de propriedades hidráulicas do solo por ensaios de evaporação em laboratório / An improved methodology for obtaining soil hydraulic properties by laboratory evaporation experiments / Leonardo Inforsato; orientador Quirijn de Jong van Lier. - - Piracicaba, 2023.

76 p.

Tese (Doutorado – Programa de Pós-Graduação em Ciências. Área de Concentração: Energia Nuclear na Agricultura e no Ambiente) – Centro de Energia Nuclear na Agricultura da Universidade de São Paulo.

1. Análise do solo 2. Condutividade hidráulica do solo 3. Evaporação 4. Física do solo 5. Fluxo dos líquidos 6. Modelos matemáticos

CDU 631.432.3 + 519.87

Elaborada por:
Marilia Ribeiro Garcia Henyei
CRB-8/3631
Resolução CFB Nº 184 de 29 de setembro de 2017

ACKNOWLEDGMENT

For this section, I prefer to use Portuguese:

Agradeço a minha família pelo apoio, Edson, Joseli e Lígia, fundamental para a realização deste trabalho.

Sou grato a Laura ♥, por me encorajar e me incentivar sempre.

Agradeço aos amigos de longa data (Fernandes, Lee, Fabio, Rui, Bigode, João, Pacheco e todos os outros, a lista é longa demais pra colocar todos aqui!) pelo apoio e descontração. Aliás, algumas conversas de bar foram utilizadas para melhorar artigos científicos já publicados. Em especial gostaria de agradecer o meu amigo Deboni, que foi quem me apresentou o CENA e o Prof. Quirijn.

Agradeço a orientação, auxílio (e paciência) do Prof. Quirijn, aprendi muito, e com certeza continuarei aprendendo com ele. O papel dele foi executado com demasiada plenitude, mesmo em questões correlatas que se estendem além do esperado de um orientador.

Agradeço aos amigos de laboratório, pela ajuda durante o trabalho, incluindo, Marina, Yeleine, Ali, Robinho, Turek, Everton, Victor, Livia, André, Thais, Thalita, Dudu, Erbesson, Lais, Elísia.

Durante o intercâmbio, fiz amizades e aprendi muito, sou grato a eles, principalmente ao Prof. Durner, que me aceitou me orientou na TU Braunschweig, e também Adil, Sascha, Andre, Ines, Deep, Mahyar, Birgit, Kai, Christian, Eric e o casal Denise e Marius.

Agradeço a secretaria de pós-graduação do CENA pelo auxílio e esclarecimento, sempre foram muito prestativos, a melhor secretaria que já conheci.

Agradeço as agências de fomento CAPES, CNPq e FAPESP - São Paulo Research Foundation (grant # 2020/00145-2) e BEPE (grant # 2021/10520-8), que financiaram este trabalho.

Obrigado!

ABSTRACT

INFORSATO, L. **An improved methodology for obtaining soil hydraulic properties by laboratory evaporation experiments**. 2023. 76 p. Tese (Doutorado) – Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, Piracicaba, 2023.

Among the methods for determining the soil hydraulic properties, the laboratory evaporation experiment is distinguished due to its experimental easiness and the simultaneous acquisition of the soil water retention function, $\theta(h)$, and conductivity function, $K(h)$. This thesis aims to improve the quality of the obtained parameters of the conductivity and retention functions using data measured in such evaporation experiments. It consists of three chapters (I, II, and III) with the respective specific objectives: (1) extension of the validity of the retention and conductivity models to the dry range of pressure heads; (2) improvement in the estimation of the $K(h)$ function when using the "simplified evaporation method" and (3) evaluating parameter transforms for the mathematical expressions of s-shaped retention functions improving the uncertainty of the fitted parameters. With the results obtained in chapter I, it is possible to extend the validity of tested and conceptualized models without the need to change or reassess the known hydraulic parameters. Results from chapter II allow obtaining values of the conductivity function with greater exactness, mainly in sandy textured soils, where the "simplified evaporation method" showed to be less accurate. Chapter III elaborates on the reduction of uncertainties in hydraulic parameters determined by non-linear fitting through parameter transformation of the hydraulic properties.

Keywords: Wind method. Non-linear regression. Soil hydraulic parameter measurement. Soil hydraulic parameter uncertainty.

RESUMO

INFORSATO, L. **Metodologia aperfeiçoada para a obtenção de propriedades hidráulicas do solo por ensaios de evaporação em laboratório**. 2023. 76 p. Tese (Doutorado) – Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, Piracicaba, 2023.

Dos métodos para determinação de propriedades hidráulicas do solo, o experimento de evaporação em laboratório se destaca pela facilidade experimental e pela aquisição das funções de retenção de água no solo, $\theta(h)$, e condutividade, $K(h)$, simultaneamente. Este trabalho visa melhorar a qualidade dos parâmetros obtidos das funções de condutividade e retenção usando dados medidos em experimentos de evaporação. O trabalho é constituído por três capítulos (I, II e III) com os respectivos objetivos específicos: (1) extensão da validade dos modelos de retenção e condutividade em zonas secas; (2) melhoria na estimativa de pontos da função de $K(h)$ utilizando o "*simplified evaporation method*" e (3) transformações de parâmetros nas expressões matemáticas das funções de retenção em formato de S melhorando a incerteza dos parâmetros ajustados. Com os resultados obtidos no capítulo I, é possível estender a validade de modelos testados e conceituados sem a necessidade de alterar ou reavaliar os parâmetros hidráulicos já conhecidos. O capítulo II permite obter valores da função condutividade com maior exatidão ao utilizar o método "*simplified evaporation method*", principalmente em solos de textura arenosa, onde demonstrava menor exatidão. O capítulo III estuda e permite a redução de incertezas em parâmetros hidráulicos determinados por ajustes não lineares, através de transformações matemáticas dos parâmetros das propriedades hidráulicas.

Palavras-chave: Método de Wind. Regressão não-linear. Estimativa de parâmetros hidráulicos. Incerteza dos parâmetros hidráulicos.

CONTENTS

GENERAL INTRODUCTION....11

REFERENCES (GENERAL INTRODUCTION)....12

1. CHAPTER I: An extension of water retention and conductivity functions to dryness....14

1.1. Introduction....14

1.2. Material and Methods....15

1.2.1. Groenevelt – Grant model as the base retention function....17

1.2.2. Kosugi and Van Genuchten models as the base retention function....17

1.2.3. Hydraulic Conductivity....19

1.2.4. Fitting procedure....22

1.3. Results and Discussion....23

1.4. Conclusion....26

1.5. References (Chapter I)....27

2. CHAPTER II: An improved calculation scheme for the Simplified Evaporation Method for obtaining soil hydraulic parameters....29

2.1. Introduction....29

2.2. Material and Methods....30

2.2.1. Original data evaluation with the simplified evaporation method....30

2.2.2. Improved evaluation by Peters et al. (2015)....32

2.2.3. Removing bias in the hydraulic conductivity function....32

2.2.4. Prediction of the suction at the center of the sample....33

2.2.5. Prediction of the suction gradient at the center of the sample....34

2.2.6. Prediction of the water flux across the center of the sample....37

2.2.7. Data Analysis....38

2.3. Results and Discussion....39

2.4. Conclusions and suggestions....43

2.5. References (Chapter II) 44

3. CHAPTER III: A fitting procedure for the soil water retention function with improved uncertainty analysis 46

3.1. Introduction 46

3.2. Material and Methods 49

3.2.1. Transforming the retention function expression 49

3.2.2. Parameter transformations 50

3.2.3. Software description 52

3.2.4. Complex-step differentiation approximation 53

3.2.5. Parameter uncertainty analysis and comparison 54

- Developed scenarios: criteria and analysis 54
- Bootstrap method 55
- Normality test 55

3.3. Results and Discussion 55

3.3.1. Description of the generated soils 55

3.3.2. Arbitrary anchoring point analysis 56

3.3.3. Anchoring point and parameter uncertainty 57

3.3.4. Parameter normality test analysis 61

- Transforms of parameters α and n 61
- Retention function fitting analysis with simultaneous transform of all parameters 66

3.3.5. Correlation analysis 71

3.4. Conclusion 72

3.5. References (Chapter III) 73

4. CONCLUDING REMARKS 76

GENERAL INTRODUCTION

The assessment and parameterization of soil hydraulic properties are necessary for any quantitative study involving the dynamics of water in the environment. Of fundamental importance for the hydrological cycle, the soil is the source from which terrestrial plants receive almost all the water necessary for their metabolism and transpiration. Currently, the ecosystem services provided by the soil are highlighted in scientific journals (Pereira et al., 2017).

For the study and quantification of hydrological processes at a soil water balance scale, two mathematical approaches stand out. In the first place, the so-called *bucket models*, such as AquaCrop developed by FAO (Salman et al., 2021). Second, models based on the Richards differential equation can be used, requiring greater knowledge of soil hydraulic properties. The Richards equation combines the mass conservation law with the Darcy-Buckingham equation.

The Richards equation allows for predicting water movement and the water content in time and space in the soil, but analytical solutions are available for only a few boundary conditions (Kool et al., 1985). Numerical solutions from robust numerical calculation procedures allow their application in many scenarios.

For the Richards equation to be solved numerically and to produce reliable output, hydraulic properties (here considered as the water retention function, RF, and hydraulic conductivity function, CF) are expressed in functions such as Burdine, 1953; Groenevelt and Grant, 2004; Kosugi, 1996; Mualem, 1976; Van Genuchten, 1980. The accurate fitting of their parameters is required to minimize prediction errors in the water balance components and soil water dynamics. Not all model parameters can be obtained through measurements, but experimental setups are available to generate data and to parameterize the RF and CF (Durner et al. 1999). The evaporation experiment (EE) introduced by Wind (1966) is such a laboratory procedure. It consists of taking pressure head measurements in a soil sample, starting at saturation and evaporating to a controlled atmosphere. The EE allows fitting of both RF and CF simultaneously.

The general objective of this thesis is to improve the accuracy of the prediction of the RF and the CF. The objective is achieved in three chapters. It extends from

considerations on commonly used models, then through improving data handling for the RF and CF and concluding in improvements in the model fitting to the soil hydraulics properties (SHP) data.

In Chapter I, an alteration in the SHP model structure, applicable to most of the commonly used models is proposed to increase their validity from the very wet (saturated) range to the oven-dry range. The modification does not require to re-fit the parameters, and, with some assumptions, does not require new data or measurements.

In Chapter II, a modification to the “simplified evaporation method” is introduced, a methodology that treats the data obtained by the evaporation experiment (proposed originally by Wind, 1966) and fits a model to these data. The assumptions for this proposed method reduce some bias implicit in the original method, significantly increasing the accuracy of the CF data for light-textured soils and increasing the accuracy for other texture classes to a lesser extent.

In Chapter III, mathematical transformations for the SHP models are introduced. These transformations facilitate the non-linear fitting procedure and approximate the uncertainty associated with these parameters to a normal distribution. The method combines well with stochastic sampling as sometimes used to generate sets of parameters for an enhanced analysis of hydrological simulations.

REFERENCES (GENERAL INTRODUCTION)

Burdine, N. T. (1953). Relative permeability calculations from pore size distribution data. *Journal of Petroleum Technology*, *5*(3), 71–78. https://doi.org/10.2118/225-G

Durner, W., Schultze, B., & Zurmühl, T. (1999). *State-of-the-art in inverse modeling of inflow/outflow experiments*. 661–681. http://www.soil.tu-bs.de/mitarbeiter/durner/public/state-of-the-art99.pdf

Groenevelt, P. H., & Grant, C. D. (2004). A new model for the soil-water retention curve that solves the problem of residual water contents. *European Journal of Soil Science*, *55*(3), 479–485. https://doi.org/10.1111/j.1365-2389.2004.00617.x

Kool, J. B., Parker, J. C., & Van Genuchten, M. T. (1985). Determining Soil Hydraulic Properties from One-step Outflow Experiments by Parameter Estimation: I. Theory and Numerical Studies 1. *Soil Sci Soc Am J*, *49*, 1348–1354.

Kosugi, K. (1996). Lognormal Distribution Model for Unsaturated Soil Hydraulic Properties. *Water Resources Research*, *32*(9), 2697–2703. https://doi.org/10.1029/96WR01776

Mualem, Y. (1976). A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research*, *12*(3), 513–522. https://doi.org/10.1029/WR012i003p00513

Pereira, P., Bogunovic, I., Munoz-Rojas, M., & Brevik, E. C. (2017). Soil ecosystem services, sustainability, valuation and management. *Current Opinion in Environmental Science & Health*, *5*, 7–13. https://doi.org/10.1016/j.coesh.2017.12.003

Salman, M., García-Vila, M., Fereres, E., Raes, D., & Steduto, P. (2021). The AquaCrop model – Enhancing crop water productivity (Vol. 47). FAO Water Report. https://doi.org/10.4060/cb7392en

Van Genuchten, M. Th. (1980). A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils. *Soil Science Society of America Journal*, *44*(5), 892–892. https://doi.org/10.2136/sssaj1980.03615995004400050002x

Wind, G. P. (1966). Capillary conductivity data estimated by a simple method. *Water in the unsatured zone*, 181–191.

1. CHAPTER I: An extension of water retention and conductivity functions to dryness

Abstract

Water retention and hydraulic conductivity are essential properties to predict water flow in soils. Commonly, these soil physical properties are represented by equations relating suction, soil water content and hydraulic conductivity. The most common empirical equations used for this purpose have a limited range of reliability, not functioning properly in the very dry range. An approach to extend the reliability and applicability of these models has been presented, making use of smoothed piecewise equations added to the base model and implying in changes in the base model parameters. We propose a modification of the model commonly known as PDI (Peters-Durner-Iden), allowing to extend the reliability of most common base models without the need to change the original parameters. The transformation is analytically equivalent, interchangeable and allows to anchor the retention function at any point instead of the residual water content. Van Genuchten, Kosugi, Brooks-Corey and Groenevelt equations can easily be combined to the proposed model, which may be used to predict water flow in the dry range where most common equations are not reliable.

1.1. Introduction

Soil water retention (RF) and hydraulic conductivity (CF) are the fundamental properties to predict water flow in soils and its linkage to key processes in the vadose zone, such as root water uptake (Dos Santos et al., 2017) and the fate of solutes and pollutants (Šimunek et al., 2018). These soil physical properties are usually represented by equations relating suction (absolute value of the pressure head), represented by h, soil water content, θ, and hydraulic conductivity, K. The most commonly used equations are those proposed by Brooks and Corey (1964), Campbell (1974), Van Genuchten (1980), Durner (1994), Kosugi (1996), Groenevelt and Grant (2004), and Peters and Durner (2008), and their parameters are calibrated using observed data in laboratory or field experiments. Different approaches to extend the range of validity of these equations to drier ranges of soil water content have been

presented (Ross et al., 1991; Rossi and Nimmo, 1994; Fayer and Simmons, 1995; Khlosi et al., 2006; Zhang, 2011; Peters, 2013). The model by Zhang (2011) consists of piecewise functions maintaining the parameters of the original model. A sound semi-empirical model to do so was presented by Peters (2013, 2014), Iden and Durner (2014), and is commonly called the PDI model.

To extend the applicability of retention functions to drier ranges, the PDI model employs a piecewise smoothed retention function, and a physical-empirical function to extend the conductivity equation. The Kosugi (Kosugi, 1994; Kosugi, 1996) and Van Genuchten – Mualem (Van Genuchten, 1980; Mualem, 1976a) (VGM), equations are the base functions used in the PDI model, though its concept is applicable to a variety of $h(\theta)$ and $K(\theta)$ functions.

To be applied, the PDI model requires a rescaling of the relative saturation function and an extension of the conductivity function, which causes its parameters to be not interchangeable with known parameters of Kosugi or VGM models. One of the reasons for parameters to change is the prediction of a non-zero value of relative saturation at the suction corresponding to oven-dry conditions by Kosugi or VGM models.

In this study we develop a modification to the PDI model allowing the use of the original Kosugi or VGM parameters, eliminating the residual water content from the equations and assuming a zero water content at the oven-dry suction. Solutions will be presented for VGM, Kosugi and Groenevelt – Grant (2004) base models.

1.2. Material and Methods

Most equations to describe soil water retention follow a structure including a residual water content θ_r [m^3 m^{-3}] and a saturated water content θ_s [m^3 m^{-3}], together with a relative saturation $S(h)$ defined as function of the suction h [m] (absolute values for the matric potential). This is the case, e.g., for the equations proposed by Brook and Corey (1964), Van Genuchten (1980) and Kosugi (1996). For these equations,

$$\theta(h) = \theta_r + (\theta_s - \theta_r) S(h) \qquad [1]$$

where the $\theta(h)$ is the water content [m^3 m^{-3}] at suction h.

In the PDI model (Peters 2013, 2014; Iden and Durner, 2014), $S(h)$ from Eq. 1 is replaced by a term $S_{cap}(h)$, the relative saturation of capillary water, which is a rescaled relative saturation from a base model (e.g. Van Genuchten, 1980 or Kosugi, 1996). Furthermore, a term $S_{ads}(h)$ is added, representing the relative saturation of the adsorbed water. The resulting $\theta(h)$ equation is:

$$\theta(h) = (\theta_s - \theta_r) S_{cap}(h) + \theta_r S_{ads}(h) \quad [2]$$

Iden and Durner (2014) defined $S_{ads}(h)$ as a smoothed piecewise function obtained by using specific mathematical techniques (Kavetski and Kuczera, 2007). Instead of using a rescaled saturation function, the proposed model applies the original RF base model, $\Gamma_{cap}(h)$, adding to it a modified adsorption component, $\theta_0 \Gamma_{ads}(h)$, resulting in

$$\theta(h) = \Gamma_{cap}(h) + \theta_0 \Gamma_{ads}(h) \quad [3]$$

in which θ_0 is the water content at the oven dry suction (h_0) of the selected base function, with Γ_{ads} (h) defined by

$$\Gamma_{ads}(h) = \frac{1}{\log \frac{h_a}{h_0}} \left\{ \log \frac{h}{h_a} + b \ln \left[1 + \exp \left(\frac{\log \frac{h_a}{h}}{b} \right) \right] \right\} \quad [4]$$

where h_0 [m] is a suction at which $\Gamma_{ads} \approx -1$, h_a [m] is a suction corresponding to an arbitrarily defined threshold value where retention switches from predominantly capillary to adsorption driven, and b [−] is a smoothing parameter, analogous to Iden and Durner (2014). Eq. 3 is further illustrated in Figure 1, clearly showing that the extended model reduces to the base model at high water contents.

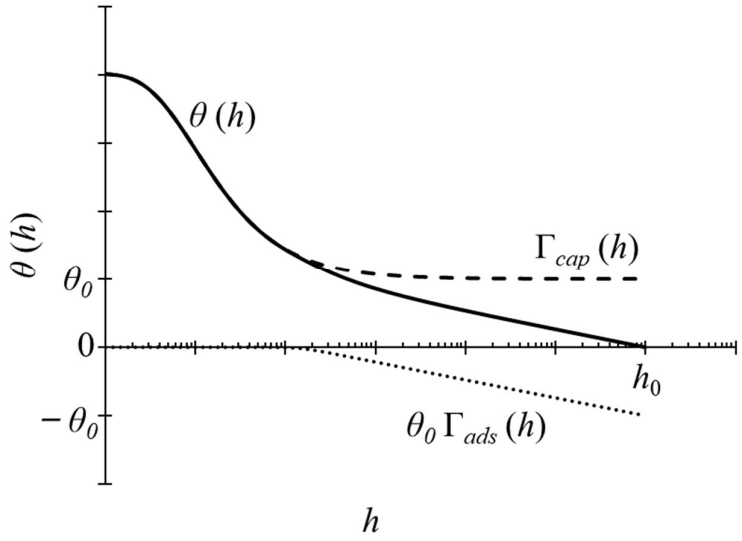

Figure 1. Graphical representation of the proposed model for the soil water retention function $\theta(h)$ (Eq. 3) composed of $\Gamma_{cap}(h)$ and $\theta_0 \Gamma_{ads}(h)$

1.2.1. Groenevelt – Grant model as the base retention function

The Groenevelt – Grant (2004) model (GG) does not include a residual water content in its formulation, and can therefore be directly combined to Eq. 3 as:

$$\theta_{GG}(h) = \theta_0 + k_1 \left\{ \exp\left(\frac{-k_0}{|h_0|^p} \right) - \exp\left(\frac{-k_0}{|h|^p} \right) \right\} + \frac{\theta_0}{\log \frac{h_a}{h_0}} \left\{ \log \frac{h}{h_a} + b \ln \left[1 + \exp\left(\frac{\log \frac{h_a}{h}}{b} \right) \right] \right\} \quad [5]$$

where $\theta_{GG}(h)$ is the water content of the GG model as function of the suction, h_0 is the suction [m] corresponding to oven-dry conditions, making $\theta_{GG}(h_0) = 0$, k_0 [m^p] is a fitting parameter, and p and k_1 are dimensionless fitting parameters. The base GG model without considering the right part of the equation including $\Gamma_{ads}(h)$ would yield $\theta(h_0) = \theta_0$.

1.2.2. Kosugi and Van Genuchten models as the base retention function

Both the Kosugi and Van Genuchten retention models include a residual water content in their formulation. Consequently, a modification needs to be performed to suit them as base functions to the proposed model. Considering a closed-form retention Eq. 1, it follows that

$$\theta(h) = \theta_r + (\theta_s - \theta_r)\int_{-\infty}^{h} f(h)\,dh = \theta_r + (\theta_s - \theta_r)\left[\int_{-\infty}^{h_i} f(h)\,dh + \int_{h_i}^{h} f(h)\,dh\right] \quad [6]$$

where $f(h)$ is a function with a known indefinite integral. Integrating Eq. 6 yields

$$\theta(h) = \theta_r + (\theta_s - \theta_r)(S_i - S_\infty) + (\theta_s - \theta_r)(S(h) - S_i) \quad [7]$$

where S_i is $S(h_i)$, and S_∞ tends to zero. Now considering the water content at an arbitrary point (θ_i, h_i) on the water retention equation

$$\theta_i = \theta_r + (\theta_s - \theta_r)S_i \quad [8]$$

and substituting Eq. 8 into Eq. 7, we obtain

$$\theta(h) = \theta_i + (\theta_s - \theta_r)(S(h) - S_i) \quad [9]$$

Isolating θ_r in Equation 8 and substituting it in Eq. 7, Eq. 9 can be written as

$$\theta(h) = \theta_i + \left(\frac{\theta_s - \theta_i}{1 - S_i}\right)(S(h) - S_i) \quad [10]$$

Eq. 10 is a $\theta(h)$ equation anchored to an arbitrary data pair (θ_i, h_i). If θ_i tends to θ_r, Eq. 10 reduces to Eq. 1. Using Van Genuchten or Kosugi base equations in the format of Eq. 10 allows to define the $\theta(h)$ equation using any arbitrary data pair (θ_i, h_i) instead of θ_r. With this approach, the equation parameters from the original base equations and the newly proposed equations are identical because the functions are analytically the same. To the best of our knowledge, Eq. 10 has not been presented before.

Using VGM in Eq 10, substituting the relative saturation $S(h)$ and S_i, the new analogous model anchored to any data pair instead of θ_r becomes

$$\theta_{VGM}(h) = \theta_i + (\theta_s - \theta_i)\left(\frac{S(h) - S(h_i)}{1 - S(h_i)}\right) \quad [11]$$

with

$$S(h) = \left[1 + (\alpha|h|)^n\right]^{\frac{1}{n}-1} \quad [12]$$

where α [m^{-1}] and n [−] a re fitting parameters.

A convenient choice for the anchor point (θ_i, h_i) could be the oven dry condition (θ_0, h_0). In this case, Eq. 10 added to Eq. 4 multiplied by θ_0 yields a new equation, analogous to the PDI model, reaching a zero water content at h_0 without further changes in the base equation (Eq. 3). This extended Eq. 10 equals

$$\theta(h) = \theta_0 + \left(\frac{\theta_s - \theta_0}{1 - S_0}\right)\left(S(h) - S_0\right) + \frac{\theta_0}{\log\frac{h_a}{h_0}}\left\{\log\frac{h}{h_a} + b\ln\left[1 + \exp\left(\frac{\log\frac{h_a}{h}}{b}\right)\right]\right\} \quad [13]$$

Equation 13 is Eq. 3 using base models in the form of Eq. 1, which may be used with VGM or Kosugi. To use VGM as base model, the relative saturation considered in Eq. 13 should agree to Eq. 12. To use the two-parameter Kosugi (1996) model as base model, $S(h)$ in Eq. 13 should be:

$$S(h) = \frac{1}{2}\operatorname{erfc}\left(\frac{\ln(h/h_m)}{\sqrt{2}\,\sigma^*}\right) \quad [14]$$

where h_m is the suction at median pore radius, erfc is the complementary error function, and σ^* [–] is a fitting parameter.

1.2.3. Hydraulic Conductivity

The capillary conductivity function is generally related to the retention function through capillary bundle models. Since the RF equation and its parameters are the same as the base equation, the CF parameters of the base model can be maintained. To extend the applicability of the CF, we used the conductivity function proposed by Peters (2013), and Peters and Durner (2008), which relates the conductivity to the adsorptive function $\Gamma_{ads}(h)$ as:

$$K_{ads} = K_{sat}\omega\left(h_0/h_a\right)^{a^*\Gamma_{ads}(h)} \quad [15]$$

where a^* [–] is a parameter related to the slope of the function in log-log scale, ω [–] is a weighting parameter, K_{sat} [m s^{-1}] is the saturated conductivity and K_{ads} [m s^{-1}] is the conductivity related to the adsorption component of the retention function.

The conductivity function related to the saturation of capillary water for VGM (Van Genuchten, 1980) is:

$$K_{cap}(h) = K_{sat} S(h)^{\lambda} \left[1 - \left(1 - S(h)^{n/(n-1)} \right)^{1-(1/n)} \right]^2 \quad [16]$$

with *S(h)* defined by Eq. 12, λ [–] is a fitting parameter and $K_{cap}(h)$ [m s^{-1}] is the conductivity related to the capillary component.

The capillary conductivity for the Kosugi (1999) model is:

$$K_{cap}(h) = K_{sat} S(h)^{\lambda} \left\{ \frac{1}{2} \operatorname{erfc} \left[\operatorname{erfc}^{-1}\left(2S(h)\right) + \frac{\sigma}{\sqrt{2}} \right] \right\}^2 \quad [17]$$

with *S(h)* defined by Eq. 14 and erfc^{-1} is the inverse of the complementary error function.

According to Peters (2013), liquid conductivity K_{liquid} can be considered the sum of capillary (K_{cap}) and adsorption (K_{ads}) components:

$$K_{liquid} = K_{cap} + K_{ads} \quad [18]$$

Generally, the conductivity value at oven dry conditions is negligible (Peters, 2014), but nevertheless, a correction to the capillary conductivity equation is proposed, according to:

$$K^*_{cap}(h) = K_{cap}(h) - K_{cap}(h_0) \left(\frac{\Omega(h) - 1}{\Omega(h_0) - 1} \right) \quad [19]$$

with

$$\Omega(h) = \left(\frac{h_0}{h_a} \right)^{\Gamma_{ads}(h)}$$

where K^*_{cap} [m s^{-1}] is the corrected capillary conductivity. Eq. 19 will gradually reduce the impact of the “residual capillary flow” at h_0, so the capillary conductivity becomes zero at $K_{cap}(h_0)$. The correction will become relevant at suctions near h_a. K_{cap} in Eq. 18 should be substituted by K^*_{cap} from Eq. 19 to make use of the proposed correction.

The most relevant equations for the proposed modified model are summarized in Table 1.

Table 1. Summary of equations for the proposed model.

Description	Equations
Conversion between residual water content θ_r and an arbitrary water content θ_i corresponding to h_i	$\theta_i = \theta_r + (\theta_s - \theta_r) S_i$ $\theta_r = \dfrac{\theta_i - \theta_s S_i}{1 - S_i}$
General retention model	$\theta(h) = \Gamma_{cap}(h) + \theta_0 \Gamma_{ads}(h)$ $\Gamma_{ads}(h) = \dfrac{1}{\log \frac{h_a}{h_0}} \left\{ \log \dfrac{h}{h_a} + b \ln \left[1 + \exp \left(\dfrac{\log \frac{h_a}{h}}{b} \right) \right] \right\}$
RF for van Genuchten – Mualem Γ_{cap}	$\Gamma_{cap}(h) = \theta_0 + \left(\dfrac{\theta_s - \theta_0}{1 - S_0} \right) \left(S(h) - S_0 \right)$ $with \quad S(h) = \left[1 + \left(\alpha \lvert h \rvert \right)^n \right]^{(1/n)-1}$
RF for Kosugi Γ_{cap}	$\Gamma_{cap}(h) = \theta_0 + \left(\dfrac{\theta_s - \theta_0}{1 - S_0} \right) \left(S(h) - S_0 \right)$ $with \quad S(h) = \dfrac{1}{2} \operatorname{erfc} \left(\dfrac{\ln(h / h_m)}{\sqrt{2}\, \sigma^*} \right)$
General conductivity model	$K_{ads} = K_{sat} \omega \left(h_0 / h_a \right)^{a^* \Gamma_{ads}(h)}$
CF for van Genuchten – Mualem	$K_{cap}(h) = K_{sat} S(h)^{\lambda} \left[1 - \left(1 - S(h)^{n/(n-1)} \right)^{1-(1/n)} \right]^2$ $with \quad S(h) = \left[1 + \left(\alpha \lvert h \rvert \right)^n \right]^{(1/n)-1}$ $K^*_{cap}(h) = K_{cap}(h) - K_{cap}(h_0) \left(\dfrac{\Omega(h) - 1}{\Omega(h_0) - 1} \right)$ $with \quad \Omega(h) = \left(\dfrac{h_0}{h_a} \right)^{\Gamma_{ads}(h)}$
CF for Kosugi – Mualem	$K_{cap}(h) = K_{sat} S(h)^{\lambda} \left\{ \dfrac{1}{2} \operatorname{erfc} \left[\operatorname{erfc}^{-1}(2S(h)) + \dfrac{\sigma}{\sqrt{2}} \right] \right\}^2$ $with \quad S(h) = \dfrac{1}{2} \operatorname{erfc} \left(\dfrac{\ln(h / h_m)}{\sqrt{2}\, \sigma^*} \right)$ $K^*_{cap}(h) = K_{cap}(h) - K_{cap}(h_0) \left(\dfrac{\Omega(h) - 1}{\Omega(h_0) - 1} \right)$ $with \quad \Omega(h) = \left(\dfrac{h_0}{h_a} \right)^{\Gamma_{ads}(h)}$

1.2.4. Fitting procedure

All RF and CF parameters can be simultaneously fitted in the proposed version of the PDI model using a standard fitting procedure. In order to show that a previously fitted base model, extended while maintaining the original parameters, leads to a good quality of fitting, a different calibration procedure is shown in the following.

We used $h(\theta)$ data pairs obtained during desorption of three soils representing different textures (Adelanto loam, Pachappa loam and Rehovot sand) retrieved from the dataset presented by Mualem (1976b). The measured data range from near-saturated (measured suction 0.1 m) to very dry conditions ($4.65 \cdot 10^4$ m for Adelanto loam, $3.19 \cdot 10^4$ m for Pachappa loam and $6.00 \cdot 10^2$ m for the Rehovot sand). The VGM analytical function was used as the base model for this fitting example.

First, the base model (VGM) was fitted to the available retention and conductivity data of the predominantly capillary part. To do so, a criterion is needed to establish the suction at which adsorption becomes most important. For the Adelanto and Pachappa loams, we used the traditional value of h = 150 m. For the Rehovot sand soil with a sand content of 97%, we searched for the best fit, adding $h(\theta)$ and $K(h)$ values until a defined suction and performing the curve fit. This procedure led to the best fit using h = 2.5 m as threshold. It is important to mention that in this kind of well-sorted soils adsorptive processes are predominant on a large range. Fitting the VGM model to data from these soils commonly yields high n and zero or slightly negative θ_r values. These parameters, however, are not meaningful, as fitting hydraulic data from such soils to the capillary VGM model does not make sense.

Following Iden et al. (2019), to minimize the weighted-least-squares objective function, a relative weight of 1000 was attributed to the water retention data and a weight of 1 to the hydraulic conductivity data. These fits yielded α and θ_r for the retention function, λ for the conductivity function, and n for both functions (retention and hydraulic conductivity). θ_s was treated as a known (measured) parameter. This fitting procedure is in line with common practice, where water contents are measured down to h = 150 m (usually in a pressure plate apparatus), and is similar to the approach assumed by Zhang (2011).

In a second step the other parameters were fitted: ω and a for CF and h_a for RF and CF, now considering all data, the entire range of suctions. K as a function of θ was

available for the Adelanto and Pachappa loams, therefore first the parameter related to the adsorption retention function, h_a, was fitted. Subsequently, water contents from $K(\theta)$ data were transformed to suctions. Then, the resulting $K(h)$ data were used to fit a and ω. For the Rehovot sand soil, K as a function of h was available and parameters ω, a and h_a were fitted simultaneously. For h_0 we assumed $6.3 \cdot 10^4$ m, as discussed in Peters (2013). For parameter b the lower limit value 0.1 from the equation proposed by Iden and Durner (2014) was used.

Another approach used to fit the proposed model was to extend the RF base model for absolute values of suctions above a critical value $|h_{crit}|$. This critical value, similar to Zhang (2011), was established using a straight line on semi-log h scale, passing through (0, h_0), the oven-dry condition, and (θ_{crit}, h_{crit}) tangent to the RF base function. The value found for h_{crit} was used for h_a.

1.3. Results and Discussion

Although the original PDI model and the here proposed version are similar, this study focused on extending the validity of already known retention and conductivity functions, maintaining the original parameters and extending reliability of the base models. On the other hand, the original PDI model focused on increasing the physical meaning and understanding of RF and CF, splitting the RF into adsorptive and capillarity components; and the CF into capillary conductivity, film conductivity and vapor conductivity. The proposed version of the retention function maintains the characteristics of differentiability and continuity and behaves linearly between adsorption and capillary components of the RF from saturation to near h_a suction, $\theta(h) = \Gamma_{cap}(h) + \Gamma_{ads}(h) \approx \Gamma_{cap}(h)$. Although it is recommended to use the adsorptive smoothed piecewise function proposed by Iden and Durner (2014), with the proposed RF version shown here the adsorption function can be easily modified to a different piecewise function or some other similar function.

Figure 2 shows the examples of fitted RF and CF using the abovementioned soils from the Mualem (1976b) database. The extended functions refer to the proposed model with VGM as base model. The $\Gamma_{ads}(h)$ component is not shown in Figure 2 because the function assumes negative values. The correction for conductivity was not used in these examples, since the value of $K(h_0)$ is of a very low order of magnitude. Parameter a^* for the adsorptive component of CF was fitted, because the value

proposed by Peters (2013) a^* = 1.5 was not satisfactory for the proposed model. The proposed model uses simple and independent functions to extend the base model, employing only three additional fitting parameters: a^* and ω for CF and h_a for RF and CF. Therefore, only some additional $h(\theta)$ and $K(h)$ data are required to extend the base model to dryness.

The threshold suction value where Γ_{ads} becomes relevant, h_a, should be considered near the limit of validity of the selected base model. Table 2 shows the fitted parameters for both analyzed soils, using VGM as the base model. Campbell and Shiozawa (1992) established a range from 10 m to 100 m as lower limit of validity for the VGM model. This agrees with results for Adelanto and Pachappa loams which fitted h_a parameters within this range. The Rehovot sand, however, showed a parameter h_a lower than this range, which agrees with results by Peters (2013). In addition, Table 3 shows the root mean square error RMSE of the fitted function relative to the real data.

Due to the linear characteristics of proposed model, if ω is already known only a few $K(h)$ data points in the adsorption region are required for fitting parameter a^*, which is related to the inclination of the hydraulic adsorption component.

Table 2. RF and CF fitting parameters for the base model VGM and for the extended model using the data from the two analyzed soils.

Soil	Fitted Parameters						
	θ_r	α [m^{-1}]	n	λ	h_a [m]	a^*	ω
Pachappa loam	0.079	$6.13\cdot10^{-1}$	2.26	0.970	40.25	12.67	$2.54\cdot10^{-7}$
Adelanto loam	0.165	$3.15\cdot10^{-1}$	2.18	0.000	36.95	10.61	$6.06\cdot10^{-5}$
Rehovot sand	0.031	4.29	3.95	0.819	0.33	49.88	$7.94\cdot10^{-7}$

Table 3. Root mean square error (RMSE) for water content θ and hydraulic conductivity, $\log_{10}(K$ [cm d^{-1}]), of the VGM base model and of the proposed extended model fitted to measured data for the two analyzed soils.

Soil	RMSE			
	θ VGM	log(K)VGM	θ Extended	log(K) Extended
Pachappa loam	0.0086	0.2518	0.0074	0.2274
Adelanto loam	0.0088	0.0942	0.0073	0.0851
Rehovot sand	0.0100	0.5920	0.0095	0.5181

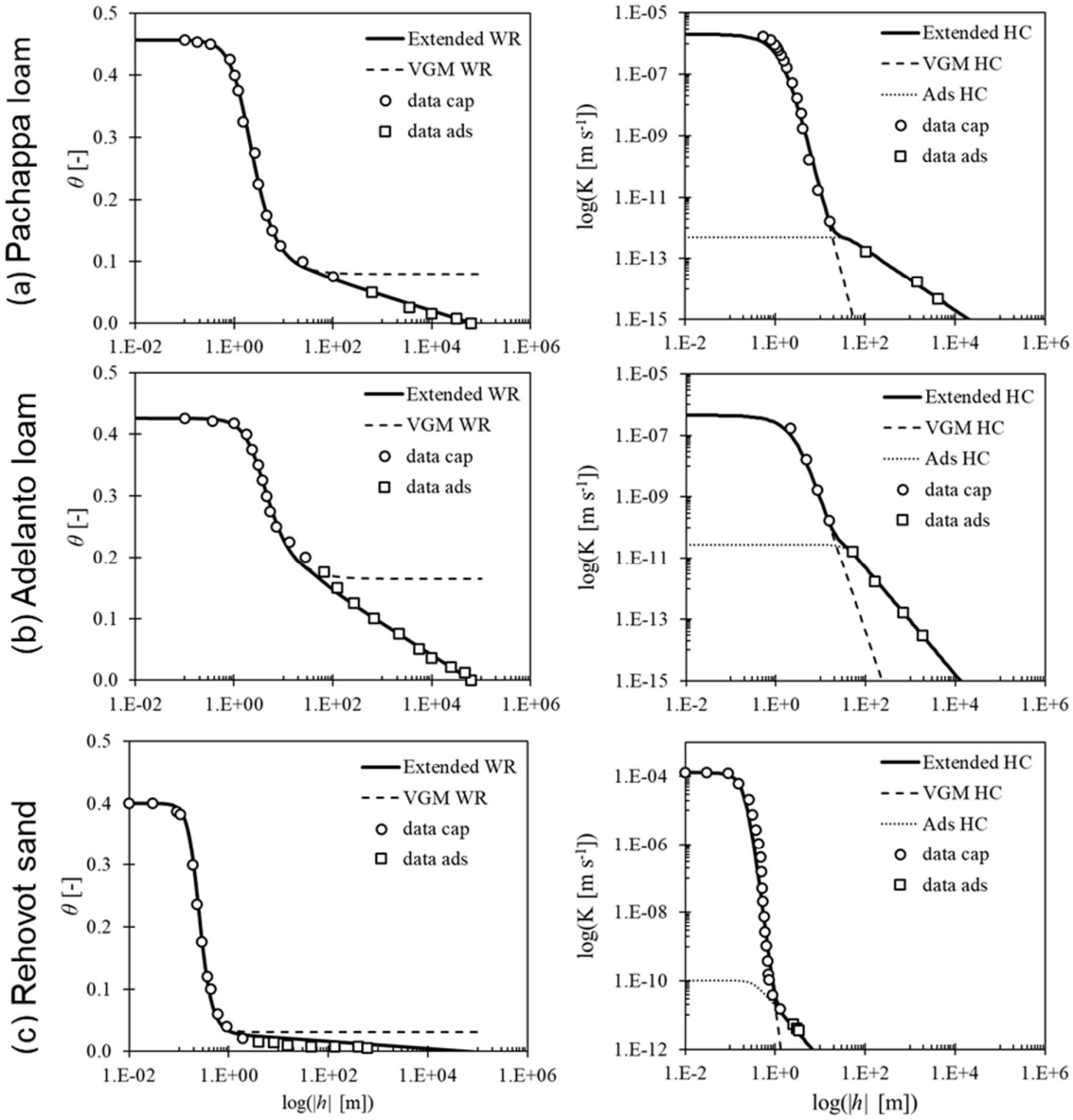

Figure 2. Water content θ and hydraulic conductivity K as function of suction h for three soils from the Mualem (1976b) database: (a) Pachappa loam, (b) Adelanto loam and (c) Rehovot sand. [Extended RF and Extended CF: fitted water retention and hydraulic conductivity functions of the extended model; VGM RF and VGM CF: fitted water retention and hydraulic conductivity functions of the base model VGM; Ads CF: fitted hydraulic conductivity function for the dryer range where the base model is not reliable; data cap: data from the capillary range ($|h| \leq 150$ m for Pachappa and Adelanto loams and $|h| \leq 2.5$ m for Rehovot sand) used to fit the base model; and data ads: data from the adsorption range ($|h| > 150$ m for Pachappa and Adelanto loams and $|h| > 2.5$ m for Rehovot sand) used to fit the remaining parameters].

Using the critical suction proposed by Zhang (2011) as h_a, its value changed to 27.55 m for the Adelanto loam and to 31.00 m for the Pachappa loam. Calculating RMSE considering all $h(\theta)$ data yielded values of $8.67 \cdot 10^{-3}$ and $7.44 \cdot 10^{-3}$ respectively. Besides the data to fit the RF of the base model and the oven-dry suction, no additional data are required to fit the critical pressure value. The Rehovot sand did not respond well to the Zhang (2011) approach, because the adsorptive component in the CF function becomes relevant at a lower absolute pressure $|h|$ than the adsorptive component of the RF function. In similar situations, parameter h_a should be fitted together with parameters ω and a^* to yield more reliable values.

The possibility of expressing the base model using any arbitrary point (θ_i, h_i) of the retention function, analytically interchangeable with the residual water content, allows to fit θ_i instead of residual water content for the base model.

1.4. Conclusion

In this study, a new version of the PDI (Peters-Durner-Iden) soil hydraulic property equations (commonly called PDI) was developed. Analyzing the results, we conclude that

1. The proposed version allows an extension of the validity range of commonly used retention and conductivity equations by summing simple functions, preserving the original parameter values.
2. The transformation is analytically equivalent, interchangeable and allows to anchor the retention function at any point, instead of residual water content.
3. The model extension is straightforward and simple, not requiring many additional dry-range data to be fitted.
4. Brooks − Corey, Van Genuchten, Kosugi and Groenevelt – Grant equations can easily be combined to the proposed model.
5. The proposed hydraulic model may be used to predict water flow in the dry range, where most common equations are not reliable.

1.5. References (Chapter I)

Brooks, R.H. and A.T. Corey. 1964. Hydraulic properties of porous media. Hydrol. Pap. Color. State Univ. Fort Collins CO 3: p. 27.

Campbell, G. S. and S. Shiozawa. 1992. Prediction of hydraulic properties of soils using particle-size distribution and bulk density data. In: M.T. van Genuchten. et al (editor), International Workshop on Indirect methods for estimating the hydraulic properties of unsaturated soils. Univ. of California, Riverside. p. 317–328.

Campbell, G.S. 1974. A simple method for determining unsaturated conductivity from moisture Retention Data. Soil Sci. 117: 311–314. doi: 10.1097/00010694-197406000-00001

Dos Santos, M.A., Q. De Jong Van Lier, J.C. Van Dam, and A.H.F. Bezerra. 2017. Benchmarking test of empirical root water uptake models. Hydrol. Earth Syst. Sci. 21: 473–493. doi: 10.5194/hess-21-473-2017

Durner, W. 1994. Hydraulic conductivity estimation for soils with heterogeneous pore structure. Water Resour. Res. 30: 211–223. doi: 10.1029/93WR02676

Fayer, M.J., and C.S. Simmons. 1995. Modified Soil Water Retention Functions for All Matric Suctions. Water Resour. Res. 31: 1233–1238. doi: 10.1029/95WR00173

Groenevelt, P.H., and C.D. Grant. 2004. A new model for the soil-water retention curve that solves the problem of residual water contents. Eur. J. Soil Sci. 55: 479–485. doi: 10.1111/j.1365-2389.2004.00617.x

Iden, S.C., J.R. Blöcher, E. Diamantopoulos, A. Peters, and W. Durner. 2019. Numerical test of the laboratory evaporation method using coupled water, vapor and heat flow modelling. J. Hydrol. 570: 574–583. doi: 10.1016/j.jhydrol.2018.12.045

Iden, S.C., and W. Durner. 2014. Comment on “Simple consistent models for water retention and hydraulic conductivity in the complete moisture range” by A. Peters. Water Resour. Res. 50: 7530–7534. doi: 10.1002/2014WR015937

Kavetski, D., and G. Kuczera. 2007. Model smoothing strategies to remove microscale discontinuities and spurious secondary optima im objective functions in hydrological calibration. Water Resour. Res. doi: 10.1029/2006WR005195

Khlosi, M., W.M. Cornelis, D. Gabriels, and G. Sin. 2006. Simple modification to describe the soil water retention curve between saturation and oven dryness. Water Resour. Res. 42. doi: 10.1029/2005WR004699

Kosugi, K. 1994. Three-parameter lognormal distribution model for soil water retention. Water Resour. Res. doi: 10.1029/93WR02931

Kosugi, K. 1996. Lognormal Distribution Model for Unsaturated Soil Hydraulic Properties. Water Resour. Res. 32: 2697–2703. doi: 10.1029/96WR01776

Kosugi, K. 1999. General Model for Unsaturated Hydraulic Conductivity for Soils with Lognormal Pore-Size Distribution. Soil Sci. Soc. Am. J. 63: 270. doi: 10.2136/sssaj1999.03615995006300020003x

Mualem, Y. 1976a. A new model for predicting the hydraulic conductivity of unsaturated porous media. Water Resour. Res. 12: 513–522. doi: 10.1029/WR012i003p00513

Mualem, Y. 1976b. A catalogue of the hydraulic properties of unsaturated soils. Res. Proj. Rep. 442: 100 p. Technion, Israel Inst. of Technol., Haifa.

Peters, A. 2013. Simple consistent models for water retention and hydraulic conductivity in the complete moisture range. Water Resour. Res. 49: 6765–6780. doi: 10.1002/wrcr.20548

Peters, A. 2014. Reply to comment by S. Iden and W. Durner on “Simple consistent models for water retention and hydraulic conductivity in the complete moisture range.” Water Resour. Res. 50: 7535–7539. doi: 10.1002/2014WR016107

Peters, A., and W. Durner. 2008. A simple model for describing hydraulic conductivity in unsaturated porous media accounting for film and capillary flow. Water Resour. Res. 44: 1–11. doi: 10.1029/2008WR007136

Ross, P.J., J. Williams, and K.L. Bristow. 1991. Equation for Extending Water-Retention Curves to Dryness. Soil Sci. Soc. Am. J. 55: 923. doi: 10.2136/sssaj1991.03615995005500040004x

Rossi, C., and J.R. Nimmo. 1994. Modeling of soil water retention from saturation to oven dryness. Water Resour. Res. 30: 701–708. doi: 10.1029/93WR03238

Šimunek, J., M.T. Van Genuchten, and R. Kodešová. 2018. Thematic Issue on HYDRUS Software Applications to Subsurface Fluid Flow and Contaminant Transport. J. Hydrol. Hydromechanics 66: 129–132. doi: 10.1515/johh-2017-0060

Van Genuchten, M.T. 1980. A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils. Soil Sci. Soc. Am. J. 44: 892. doi: 10.2136/sssaj1980.03615995004400050002x

Zhang, Z.F. 2011. Soil Water Retention and Relative Permeability for Conditions from Oven-Dry to Full Saturation. Vadose Zone J. 10: 1299–1308. doi: 10.2136/vzj2011.0019

2. CHAPTER II: An improved calculation scheme for the Simplified Evaporation Method for obtaining soil hydraulic parameters

Abstract

The Richards equation is commonly used to predict soil water dynamics and soil water storage, requiring the fitting of the water retention function (RF) and the hydraulic conductivity function (CF). The method based on a laboratory evaporation experiment (EE) followed by a fitting procedure is a traditional technique to fit RF and CF (together referred to as soil hydraulic properties, SHP) for a specific soil. The “simplified evaporation method” makes considerations to simplify the numerical procedure of the SHP acquisition. We modified the “simplified evaporation method” to provide more precise data for the SHP fitting. The results show more accurate data for the conductivity function for different soil texture ranges, including for sandy soils for which the original method produced data with undesirable bias.

2.1. Introduction

Accurate soil hydraulic properties (SHP) are required for the prediction of soil water dynamics. A commonly used procedure to estimate the SHP for a specific soil is the Simplified Evaporation Method (SEM), first presented by Schindler (1980) and later improved by Peters and Durner (2008) and Peters et al. (2015). The SEM consists of a laboratory experimental step followed by a computational step (data evaluation and fitting procedure). The experimental step is performed in a soil sample, initially saturated with water, and subjected to evaporation while the data measurements are taken. The computational step is characterized by the evaluation of the data acquired during the laboratory experiment, using specific assumptions for the SEM, and is followed by a fitting procedure to finally obtain the SHP parameters (retention - RF and conductivity - CF functions) simultaneously (Peters et al., 2015).

The SEM assumptions allow obtaining SHP with relative simplicity, requiring simple analytical calculations for data evaluation and a non-linear fitting procedure for obtaining the parameters. The soil hydraulic parameters estimated through SEM show some systematical inaccuracy for sand soils (Iden et al., 2019) and some SHP

differences compared with inverse modeling (Dettmann et al., 2019). The differences arise from non-linearities in the suction, h [cm], (absolute value of the pressure head), as function of the vertical position ("depth") in the sample. The "original SEM" (considered as the procedure described in Peters et al., 2015) cannot predict correctly these non-linearities in the data evaluation step.

We introduce a new method for evaluating the data from evaporation experiments to parameterize the hydraulic conductivity function, $K(h)$ [cm d^{-1}]. The new method is composed of three new approaches to calculate (K, h) pair of values, changing the "original SEM" assumptions. The objective is to reduce the bias observed in the $K(h)$ of the original SEM, using new approaches to estimate conductivity data.

2.2. Material and Methods

2.2.1. Original data evaluation with the simplified evaporation method

The procedure of the evaporation experiment using the SEM is characterized by the measurement of the suction h [cm] at two vertical positions ("depths") in a soil sample using tensiometers. The average water content in the sample over time, t [d], is calculated by weighing on a balance. Initially, the soil sample is saturated with water, and evaporation occurs from its surface until the sample is relatively dry and tensiometers stop functioning. During this process, sample weight and tensiometer readings (suctions) are performed periodically. Thus, at the end of the experiment, a time series composed of average water content and both tensiometer readings are available. Although not necessary, the oven-dry weight of the soil sample may be taken at the end of the evaporation experiment as a reference.

Figure 3 illustrates the experimental setup of an evaporation experiment. The length of the column is L [cm], the average volumetric water content of the soil at a discrete time t_i is θ_i [cm^3 cm^{-3}], and the suctions measured by the two tensiometers are denoted by $h_{i,1}$ [cm] and $h_{i,2}$ [cm]. The vertical positions of the tensiometers are $z_1 = 0.25\,L$ and $z_2 = 0.75\,L$. The vertical axis z is positive upwards and $z = 0$ is at the bottom of the soil sample.

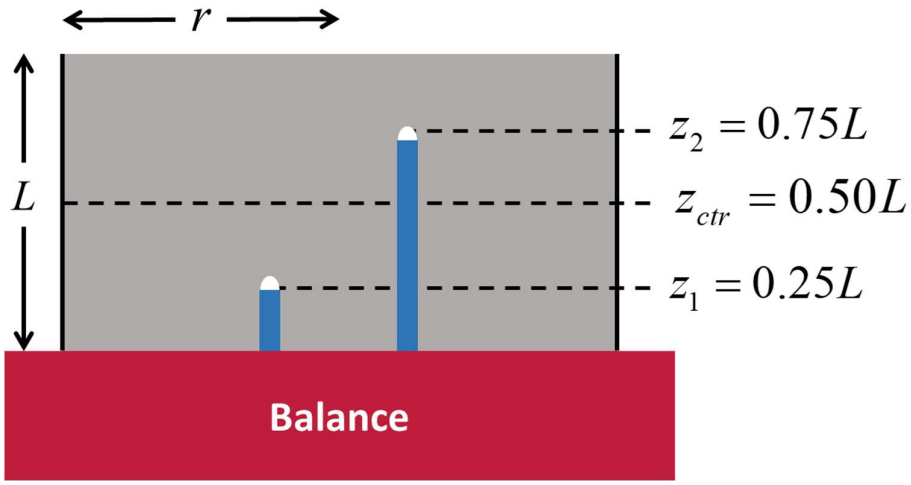

Figure 3. Profile view of the experimental setup for the evaporation experiment. The length of the column of the soil sample is L [cm], the red rectangle represents the balance, r is the radius of the soil sample [cm], the vertical position in the soil sample [cm] is represented by z (where $z = 0$ cm is taken at the sample bottom), z_2 and z_1 are the vertical positions of the upper and lower tensiometer, respectively, and z_{ctr} is the vertical position of the center of the soil sample (adapted from Peters et al., 2015).

The original SEM calculates the suction in the center of the soil sample, $\overline{h_{i,ave}}$ [cm] as the arithmetic mean of the two suctions $h_{i,1}$ and $h_{i,2}$ at time t_i (Peters and Durner, 2008; Schindler, 1980):

$$\overline{h_{i,ave}} = \frac{h_{i,1} + h_{i,2}}{2} \qquad [20]$$

The hydraulic conductivity is calculated as:

$$K_{i-\frac{1}{2}} = -\frac{q_{i-\frac{1}{2}}}{\nabla H_{i-\frac{1}{2}}} \quad with \quad \nabla H_{i-\frac{1}{2}} = \nabla h_{i-\frac{1}{2}} - 1 = \frac{\left(\overline{h_{i,ave}} - \overline{h_{i-1,ave}}\right)/2}{z_2 - z_1} - 1 \qquad [21]$$

where $\nabla H_{i-\frac{1}{2}}$ is the vertical pressure gradient at the center of the soil sample, $\nabla h_{i-\frac{1}{2}}$ is the suction gradient with respect to the soil sample depth at the center of the sample,

$q_{i-\frac{1}{2}}$ is the water flow at the center of the soil sample at time $t_{i-\frac{1}{2}}$. The water flux across the center, $q_{i-\frac{1}{2}}$, is computed as:

$$q_{i-\frac{1}{2}} = 0.5 \cdot L \frac{\overline{\theta}_i - \overline{\theta}_{i-1}}{t_i - t_{i-1}} \quad [22]$$

where $\bar{\theta}_i$ is the average water content in the soil sample estimated by the weight at time t_i.

2.2.2. Improved evaluation by Peters et al. (2015)

Peters et al., (2015) proposed improvements to the SEM process to incorporate some non-linearities in space and time, specifically changing the considerations for $\overline{h_{i,ave}}$, which becomes $\overline{h_{i,mix}}$, and is calculated as

$$\overline{h_{i,mix}} = \varphi_{avg}\,\overline{h_{i,ave}} + \left(1 - \varphi_{avg}\right)\overline{h_{i,geo}} \quad [23]$$

where $\overline{h_{i,geo}}$ is the geometric mean of the suction

$$\overline{h_{i,geo}} = \frac{\sqrt{\left(h_{i,1} + h_{i-1,1}\right)\cdot\left(h_{i,2} + h_{i-1,2}\right)}}{2} \quad [24]$$

and φ_{avg} is a weighting factor calculated as

$$\varphi_{avg} = \frac{1}{\nabla h} \quad [25]$$

For the retention function (RF), the value pairs (θ; h) are estimate from the average water content in the whole soil sample ($\overline{\theta_i}$) at time t_i, and the correlated suction $\overline{h_{i,mix}}$. So, the RF data used for fitting are the pairs ($\overline{\theta_i}$;$\overline{h_{i,mix}}$). The average water content is calculated from the weight of the soil sample.

2.2.3. Removing bias in the hydraulic conductivity function

As shown by Iden et al. (2019), the improved scheme of Peters et al. (2015) still leads to a bias in the calculated conductivity data, especially for light-textured (sandy) soils. This is caused by the fact that the spatial distribution of suction $h(z)$ is very nonlinear in many sandy-textured and peaty soils. In general, the errors in the calculated data points of the CF are caused by an error in the average suction

(abscissa of $K(h)$ function) or the calculated hydraulic conductivity (ordinate). As the latter is calculated from the hydraulic head gradient and the water flux across the center of the column, three error sources can be distinguished corresponding to the calculation of:

1. the average suction to which the conductivity is assigned (h_{ctr}),
2. the suction gradient at the center of the column (∇h_{ctr}),
3. the water flux density across the center (q_{ctr}).

In the following, we introduce three revised calculation schemes that decrease these errors and their propagation in the CF.

2.2.4. Prediction of the suction at the center of the sample

Instead of analyzing the functions $h_1(t)$ and $h_2(t)$, we start with the respective reverse functions $t_1(h)$ and $t_2(h)$ as illustrated in Figure 4. These functions return the time of arrival of a suction h at the position of the two tensiometers. First, we select a suction measured by the bottom tensiometer, h_{i1}, and calculate the average time for this suction to be reached by:

$$t_{ctr,i} = \left(\frac{t_1(h_{i,1})^w + t_2(h_{i,1})^w}{2} \right)^{1/w} \quad [26]$$

where w is a weighting factor. For $w = 1$, the average becomes arithmetic, and if $w \to 0$, the average becomes geometric. The time $t_{ctr,i}$ can be interpreted as the time at which the measured suction h_{i1} reaches the center of the soil column. This allows us to calculate the suction at the center of the column as function of time by applying the reverse function. We used $w = 2$, which yielded good results and for which the arrival time $t_{ctr,i}$ becomes greater than the arithmetic (and the geometric) means corresponding to the two times. Figure 4 illustrates the calculation.

The concept for the calculating the value of h as an average of time refers to the time each specific value of h takes to move from the upper to the lower tensiometer.

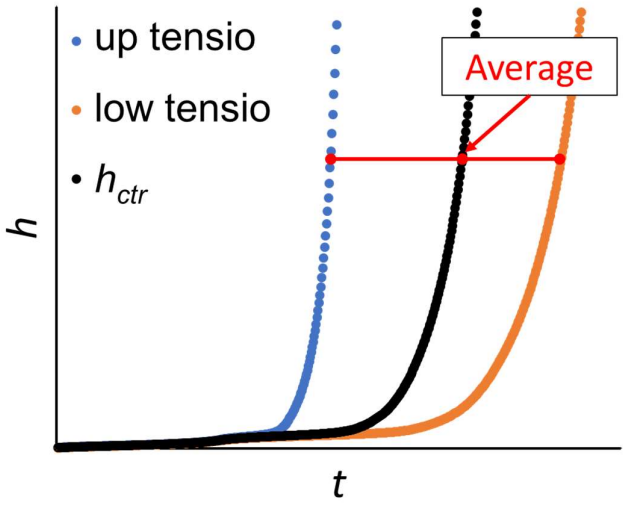

Figure 4. Time series of the measured suction (h) and the calculation of the suction as function of time (t) at the center of the soil column by Eq. (26), with upper and lower tensiometer readings (up tensio and low tensio respectively) and estimates for the average suction at the center of the sample.

2.2.5. Prediction of the suction gradient at the center of the sample

The calculation of the spatial derivative of the suction $\nabla h = \frac{dh}{dz}$ is based on the concept outlined in the preceding section. In general, the position and time of an arbitrary suction h in between the tensiometers can be calculated by it. To calculate ∇h at the center of the column, the values of the suction a little above (upward), h_{upw}, and a little below (downward), h_{dnw}, the center of the soil sample profile are required. When using the arithmetic mean to calculate $t_{ctr,i}$ (equivalent to considering the velocity of downward movement of a specific value h as constant), this procedure is intuitive, but for the average considering $w = 2$ in Eq. 26 it is less straightforward. Considering Δz^* [cm] as a small absolute distance from h_{ctr} to h_{upw} and from h_{ctr} to h_{dnw}, to calculate the respective time for h_{upw} and h_{dnw}:

$$\Delta t_{sq} = t_{top}^{\ 2} - t_{bot}^{\ 2} \qquad [27]$$

$$\Delta t_{eq} = \frac{\Delta t_{sq} \cdot \Delta z^*}{dist} \qquad [28]$$

where the t_{top} and t_{bot} are the respective value of time for the upper and lower tensiometers to reach a specific value of h, the Δt_{eq} is the squared equivalent time in

which the considered value of h is in relation to Δz^* and $dist$ is the distance between tensiometers. So, the respective times for h_{upw} and h_{dnw} can be calculated as

$$t_{upw} = \sqrt{t_{ctr}^{\,2} + \Delta t_{eq}}$$
$$t_{dnw} = \sqrt{t_{ctr}^{\,2} - \Delta t_{eq}} \qquad [29]$$

Calculating t_{upw} and t_{dnw} for each specific h value, the respective points (h_{upw}, t_{upw}) and (h_{dnw}, t_{dnw}) are obtained. Figure 5 shows the behavior of h_{ctr}, h_{upw} and h_{dnw} as a function of time, together with the measurements of the upper tensiometer, lower tensiometer, and the true values of h_{ctr} (True h_{ctr}).

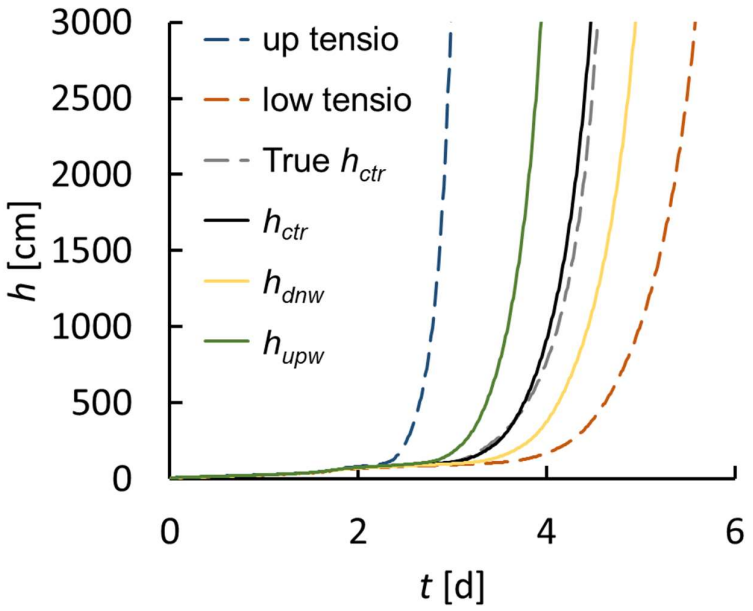

Figure 5. Suction as a function of time of the upper tensiometer (up tensio), lower tensiometer (low tensio), the true value of the suction at the center of the soil sample profile (True h_{ctr}), the estimated suction at the center of the soil sample profile (h_{ctr}) and the estimated suction at distance Δz^* above the center of the soil sample (h_{upw}), and at Δz^* below the center of the soil sample (h_{dnw}).

Finally, the vertical gradient at the center of the soil sample profile at each time is calculated as

$$\nabla h_{ctr} = \frac{h_{upw} - h_{dnw}}{2 \cdot \Delta z^*} \qquad [30]$$

where $2 \cdot \Delta z^*$ is the distance between respective suctions h_{upw} and h_{dnw} at a specific time. Δz^* must have the same value as in Eq. 28.

Fig 6 illustrates the process of obtaining h_{ctr}, h_{upw} and h_{dnw}, and therefore the gradient. From graph (1) to (2), the inversion for the tensiometer data is calculated, from (2) to (3), the values for t_{ctr} (black line), t_{upw} (green) and t_{dnw} (yellow), are calculated as averages (Eqs. 26 and 29). From (3) to (4) all inverses are undone, that is, $t(h)$ data is transformed to $h(t)$ data, returning to values of h as function of t. The gradient is calculated using h_{upw} (green) and h_{dnw} (yellow) in (4), as described in Eq. 30.

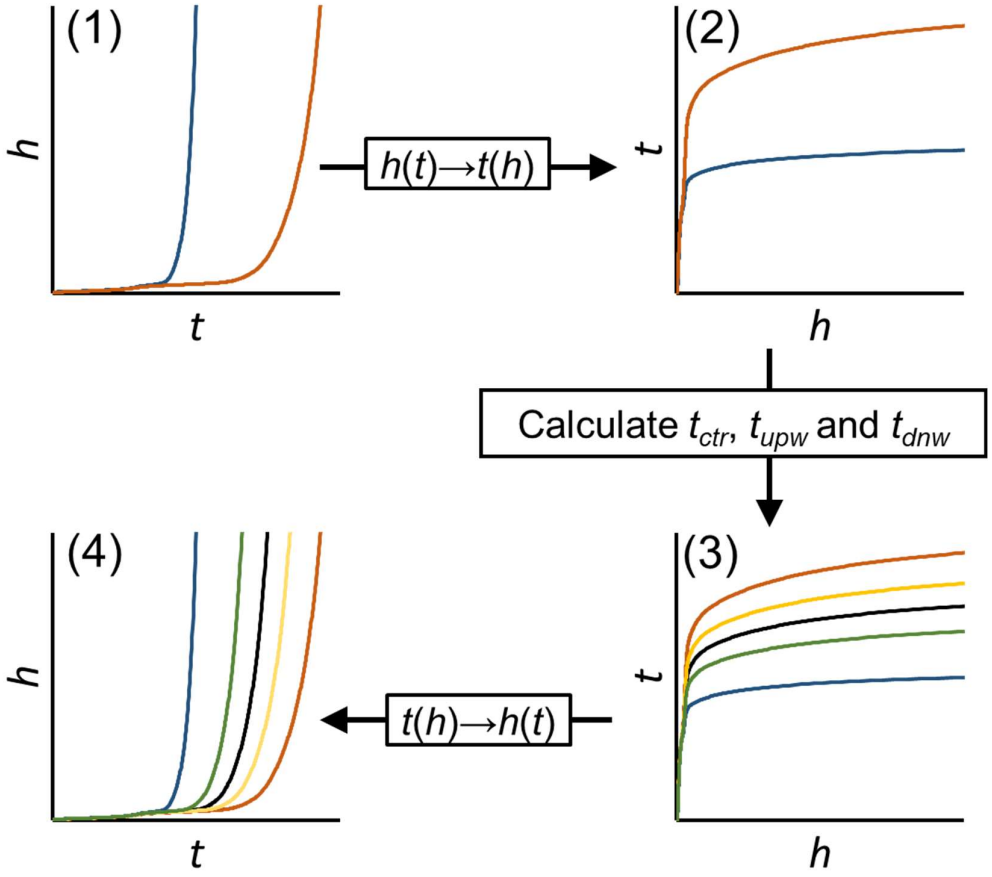

Figure 6. Process flow for obtaining h_{ctr}, h_{upw} and h_{dnw} from both tensiometers data (orange and blue lines). Graphs (1) and (4) show suction (h) as a function of time (t). Graphs (2) and (3) show t as a function of h. In Graph (3), t_{ctr} (black line), t_{upw} (green) and t_{dnw} (yellow) are represented, and in Graph (4) their respective inverse functions h_{ctr}, h_{upw} and h_{dnw}.

2.2.6. Prediction of the water flux across the center of the sample

The flux at the center of the soil sample, q_{ctr}, is calculated based on the measurements of the scale, the bottom tensiometer and the estimate of h_{ctr}. Two consecutive measurements (represented by the subscripts i-1 and i) of h_{bot} with estimates of h_{ctr} can be transformed into q_{ctr}, which is the water content flux at the center of the soil sample profile. The first step is to estimate the water content at the center of the soil sample (θ_{ctr}) and the water content related to the lower tensiometer (θ_{bot}). For this, interpolated values from the RF data ($\overline{\theta_i}; \overline{h_{i,mix}}$) are used as an approximation for $h_i \rightarrow \theta_i$ (as described in topic "Improved evaluation by Peters et al. (2015)").

$$\begin{aligned} h_{ctr,i} &\rightarrow \theta_{ctr,i} \\ h_{bot,i} &\rightarrow \theta_{bot,i} \\ h_{ctr,i-1} &\rightarrow \theta_{ctr,i-1} \\ h_{bot,i-1} &\rightarrow \theta_{bot,i-1} \end{aligned} \qquad [31]$$

Then, the water storage, l_{wc} [cm], is calculated as

$$\begin{aligned} l_{wc,i} &= \theta_{ctr,i} \cdot \left(\frac{z_{ctr} - z_1}{2} \right) + \theta_{bot,i} \cdot \left(\frac{z_{ctr} + z_1}{2} \right) \\ l_{wc,i-1} &= \theta_{ctr,i-1} \cdot \left(\frac{z_{ctr} - z_1}{2} \right) + \theta_{bot,i-1} \cdot \left(\frac{z_{ctr} + z_1}{2} \right) \end{aligned} \qquad [32]$$

where θ_{bot} is the water content related to the bottom tensiometer suction, and θ_{ctr} is the average water content at the center of the soil sample (at $z = 0.5\,L$). After calculating l_{wc} for both consecutive times, t_{i-1} and t_i, the $q_{1-\frac{1}{2}}$ at a specific time $t_{1-\frac{1}{2}}$, is calculated as

$$q_{ctr} = \frac{l_{wc,i} - l_{wc,i-1}}{\left(t_i - t_{i-1}\right)} \quad \text{at time } t_{1-\frac{1}{2}}, \text{ and } \quad t_{1-\frac{1}{2}} = \frac{t_i + t_{i-1}}{2} \qquad [33]$$

where $l_{wc,i-1}$ is the l_{wc} at t_{i-1}, and $l_{wc,i}$ is the l_{wc} at t_i. The new q_{ctr} substitutes the original $q_{1-\frac{1}{2}}$ for the estimation of K, using the data pair (q_{ctr}, $t_{1-\frac{1}{2}}$).

2.2.7. Data Analysis

To evaluate the proposed procedure, evaporation experiments were simulated using HYDRUS-1D (Šimůnek et al., 2013), which solves Richards equation by finite element formulation. For the virtual evaporation experiment, the simulations considered capillarity and adsorption for the flow of liquid water using the PDI model (Iden and Durner, 2014; Peters, 2013, 2014) coupled with isothermal vapor flow at a constant 20º C. Three different soils of a broad range of textures (Rehovot Sand, Sandy Loam, and Clay Loam) were used. Table 4 presents the parameters for each soil. All the simulations considered soil sample height $L = 5\,\mathrm{cm}$, 101 equidistant nodes, and 15 days of simulation. A constant surface flow of 1.0 cm/day or a maximum surface suction equal to 10^6 cm was used. More details of the simulations are available at Peters et al., (2015).

Table 4. Soil water parameters for the PDI model, h_m [cm] is the suction corresponding to the median pore radius, σ^* [−] is the standard deviation of the lognormal density function, θ_r [−] is the residual water content, θ_s [−] is the saturated water content, τ [−] is the tortuosity factor, *Ks* [cm d^{-1}] is the saturated water content and ω [−] is the weighting factor for the capillary and film flow.

Soil	h_m [cm]	σ^* [-]	θ_r [-]	θ_s [-]	τ [-]	*Ks* [cm d^{-1}]	ω [-]
Rehovot Sand	25.0	0.62	0.030	0.40	0.55	1700	$9.1\cdot10^{-7}$
Sandy Loam	198	1.24	0.083	0.43	-0.52	8.00	$2.5\cdot10^{-4}$
Clay Loam	442	1.36	0.295	0.50	-0.91	0.65	$2.7\cdot10^{-3}$

To investigate the proposed theory, noise was added to the simulated evaporation experiment data. Two levels of noise were added to the suction data representing the tensiometer readings. In scenario (i), a standard deviation of $\sigma_h = 0.2\,\mathrm{cm}$ was added to both tensiometer suction values (random and normally distributed), and for scenario (ii), a standard deviation $\sigma_h = 1.0\,\mathrm{cm}$ was used. For all soils and scenarios, the estimates for the CF points were analyzed and compared "New SEM" considering all three proposed modifications (h_{ctr}, ∇h_{ctr} and q_{ctr}). Since the new proposal for q_{ctr} relies on tensiometer suction readings, subjected to higher errors than the balance, a second procedure was tested, "New SEM (q/2)", considering only

h_{ctr} and ∇h_{ctr} (and the flux is estimated with the original procedure, $q_{i-\frac{1}{2}}$). The equations to calculate CF data with New SEM and New SEM (q/2) respectively become:

$$K\left(h_{ctr}\right)_{t_{i-\frac{1}{2}}} = -\frac{q_{ctr}}{\nabla h_{ctr} - 1} \quad [34]$$

$$K\left(h_{ctr}\right)_{(q/2),t_{i-\frac{1}{2}}} = -\frac{q_{i-\frac{1}{2}}}{\nabla h_{ctr} - 1} \quad [35]$$

For the comparison of the estimated SEM data, a nonlinear fitting procedure was performed to find the best conductivity parameters for the model considering the noisy data for both scenarios. Then, the RMSD between the true function was compared with the resulting fitting function for scenarios (i) and (ii) respectively for each soil. For this comparison, values between pF 1.75 and pF 5.00 (pF defined as $\log_{10}(h)$, for h in cm; pF 1.75 corresponding to 56.2 cm; pF 5.00 to 100,000 cm), were used for RMSD calculation.

2.3. Results and Discussion

Figure 7 shows the plot of the true conductivity function, the new proposed SEM data calculation, and the original SEM calculation for the Rehovot Sand soil without noise. Visually, the introduced SEM estimate provided a result closer to the true curve. The bias of the original method was already discussed in Iden et al., (2019) for sandy soils, and is caused by non-linearities of the gradient and the approximation of the flux at the center of the sample. The noisy data around pF 1.0 and pF 2.0 are due to computational numerical inaccuracies. To reduce the error in the gradient calculation, 0.2 cm was used for Δz^*, for all gradient ∇h_{ctr} calculations.

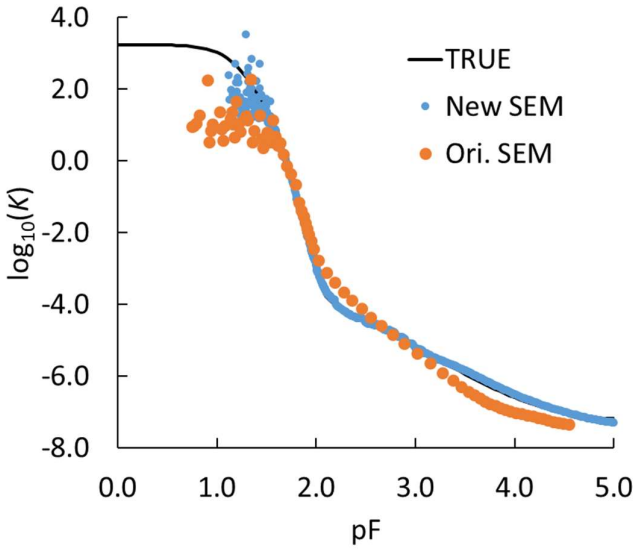

Figure 7. Logarithm of the hydraulic conductivity K [cm d^{-1}] as a function of the pF, showing the true function used for the simulation for the Rehovot Sand soil, the data estimated with the proposed (New SEM) and original (Ori. SEM) calculation procedures.

Figure 8 presents the CF and the estimated SEM data for Rehovot Sand soil with added noises, showing in the left two graphs data calculated with the new proposed SEM considering New SEM (h_{ctr}, ∇h_{ctr} and q_{ctr}) and in the right graphs data calculated with New SEM (q/2), (h_{ctr} and ∇h_{ctr}), with flux density calculated by the original procedure. The greater noise in the New SEM compared to the New SEM (q/2) is caused by the propagation of the error of the suction values, since the new method calculates q_{ctr} using the respective tensiometer reading, while the original method uses only weight data from the balance.

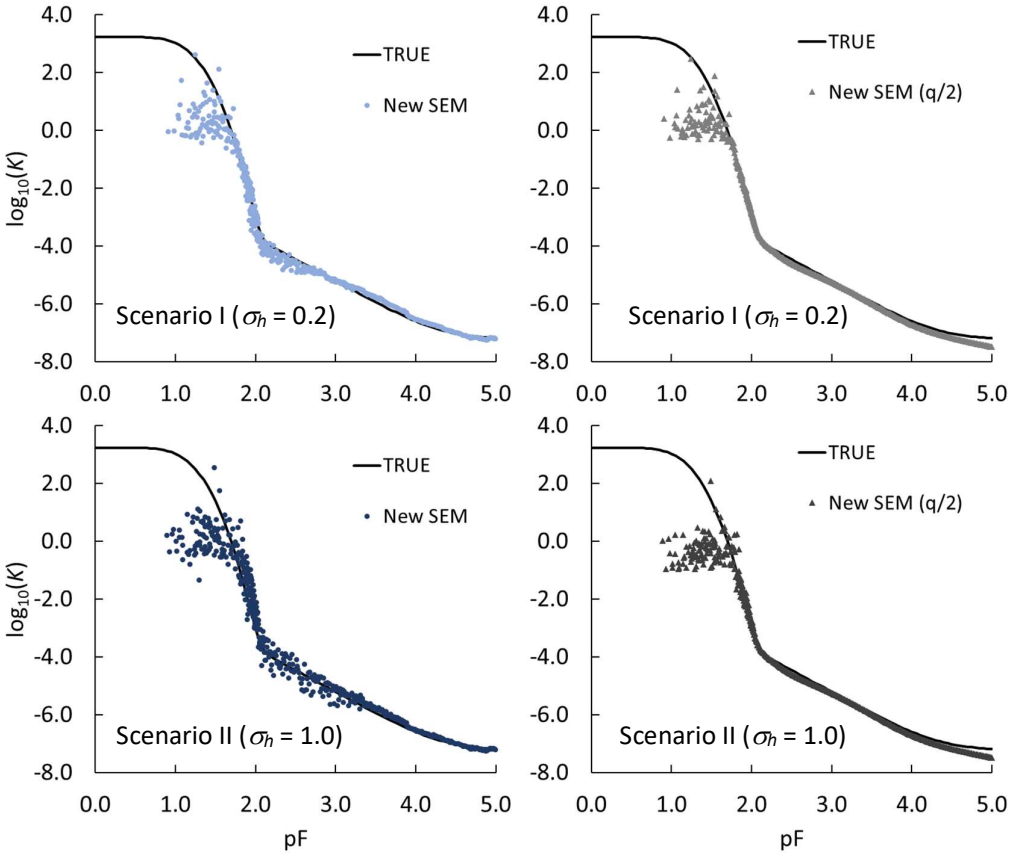

Figure 8. $\text{Log}_{10}(K$ [cm d^{-1}]) as a function of pF. The continuous line is the true conductivity. The dots and triangles are estimates provided by SEM. Top figures: Scenario I with σ_h = 0.2 cm. Bottom figures: Scenario II with σ_h = 1.0 cm. The left figures are calculated by the New SEM, the right figures with the New SEM, but considering q_{ctr} as half of the flow density at the surface.

The fitted conductivity functions of the noisy data are shown in Fig. 9. All three soils and both scenarios are displayed. For soils Sandy Loam and Clay Loam, the fitted curves are overlapping, showing neither worsening nor improvement in the fits when comparing the new method with the original method to provide the SEM data. For the Rehovot Sand soil, an improvement is observed when using the new method (New SEM fit) and the new method with the original density flux (New SEM (q/2) fit) when compared with the original SEM fit. The erroneous bias observed in Fig. 7 and by Iden et al. (2019) disappeared when using the proposed method.

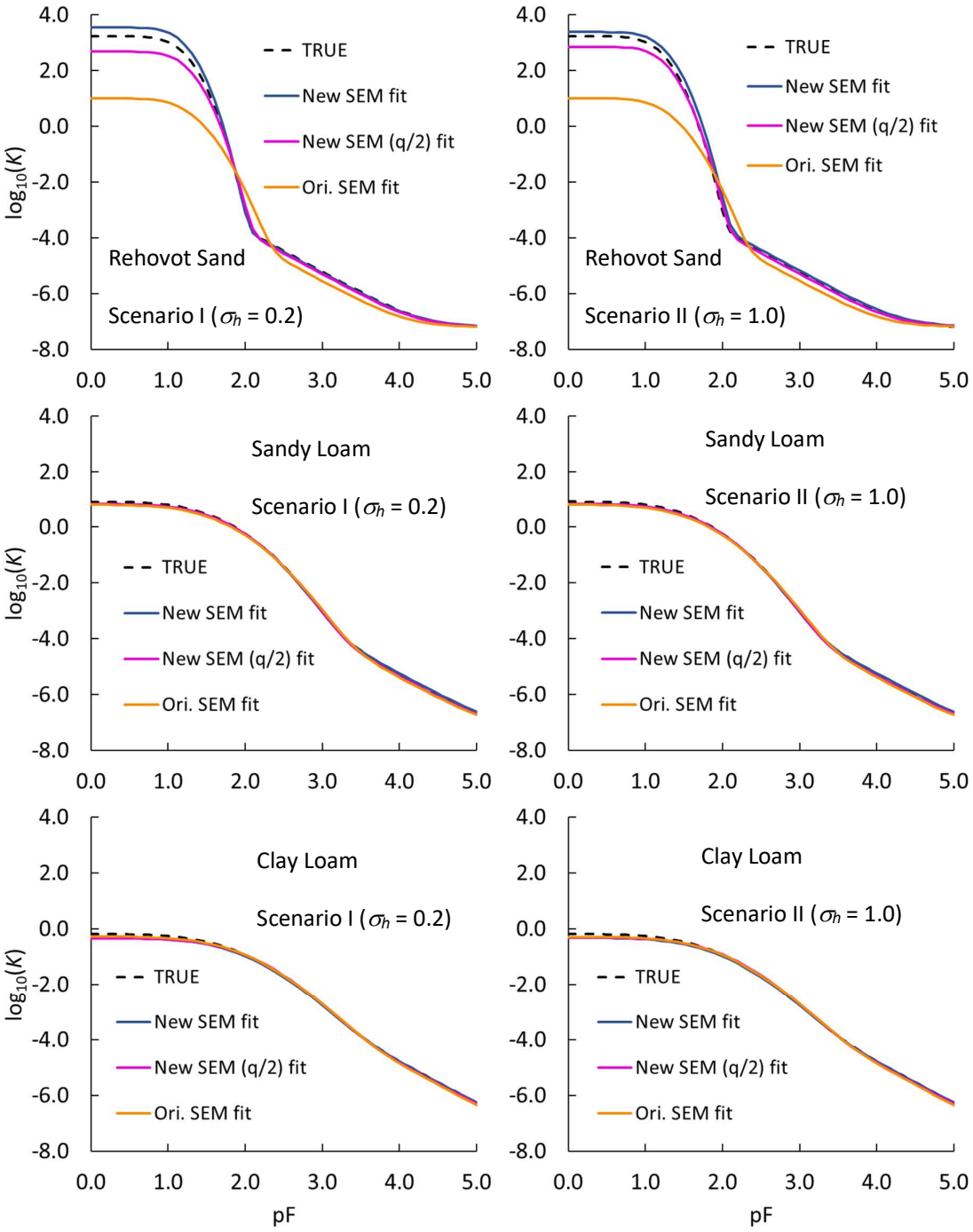

Figure 9. $\mathrm{Log}_{10}(K\,[\mathrm{cm\ d^{-1}}])$ as a function of pF of the true conductivity function (dashed black line) and the fitted functions $K(h)$ for the soils: Rehovot Sand, Sandy Loam and Clay Loam, for scenarios I (left figures, lower noise level) and II (right figures, higher noise level).

The RMSD comparing the fitted functions with the true function showed improvement when using the new method (Table 5). The new method using h_{ctr}, ∇h_{ctr}

and q_{ctr}, provided lower values of RMSD overall, except for the case of the Rehovot Sand soil with noise $\sigma_h = 1.0\ \mathrm{cm}$ where the RMSD was lower than for the new method (h_{ctr} and ∇h_{ctr}), but considering the original q_{ctr}. This is an indication that the q_{ctr} calculation using tensiometer measurements can cause a higher uncertainty than using the original flux for calculation. However, it is important to mention that $\sigma_h = 1.0\ \mathrm{cm}$ is a relatively high uncertainty for modern tensiometer readings. Anyway, even at a higher uncertainty, both proposed methods provided better results than the original procedure.

Table 5. RMSD of the logarithm values (base 10) of the fitted $K(h)$ functions compared to the true function using different added noise levels (σ_h). Original fit is the original SEM methodology, New SEM fit uses the new proposal for the SEM data for the fitting and New SEM (q/2) fit considers the new calculation for SEM but using the original flux density at the center of the soil sample.

	Noise Level	$\mathrm{Log}_{10}(K)$ RMSD		
	σ_h	New SEM fit	New SEM (q/2) fit	Original fit
Rehovot Sand	0.2	0.2775	0.3556	0.6340
	1.0	0.4716	0.3605	0.6348
Sandy Loam	0.2	0.2304	0.3267	0.3741
	1.0	0.2304	0.3267	0.3741
Clay Loam	0.2	0.2259	0.2711	0.3070
	1.0	0.2212	0.2753	0.3075

2.4. Conclusions and suggestions

We introduced a new method for parameterizing the hydraulic conductivity function $K(h)$ by the simplified evaporation method, SEM, (Schindler, 1980). The proposed method improved the SEM data for the $K(h)$ when compared to the original method proposed by Peters et al., (2015). With a statistical bias added to the tensiometer data, the new estimation also predicted more accurately the data for $K(h)$ for a broad range of soil texture scenarios simulated.

For each component used to calculate SEM $K(h)$ data, a new approach was presented. Two levels of modification were studied: $K(h)$ considering the new approach for h_{ctr}, ∇h_{ctr} and q_{ctr}, and $K(h)$ considering the new approach for h_{ctr} and ∇h_{ctr} but

keeping the original method for the flow q_{ctr}). Both levels of modification were tested for two different statistical noise levels added to the exact data. In all scenarios, the new method provided better results. The comparison in between the two classes, New SEM provided lower RMSD in almost all realized tests, only for higher noise level with soil Rehovot Sand the New SEM (q/2) provided lower RMSD.

In general, the results removed the bias that SEM estimates for $K(h)$, especially in light-textured soils. For future study, we suggest comparison with more tested methods for determining hydraulic properties, e.g. inverse modeling, and also to estimate the magnitude of the difference of the proposed method compared to the original method on actual evaporation experiment data. Also, the authors suggest testing h_{ctr} (instead of $\overline{h_{i,mix}}$) to determine the retention function through SEM.

2.5. References (Chapter II)

Dettmann, U., Bechtold, M., Viohl, T., Piayda, A., Sokolowsky, L., & Tiemeyer, B. (2019). Evaporation experiments for the determination of hydraulic properties of peat and other organic soils: An evaluation of methods based on a large dataset. Journal of Hydrology, 575, 933–944. https://doi.org/10.1016/j.jhydrol.2019.05.088

Iden, S. C., Blöcher, J. R., Diamantopoulos, E., Peters, A., & Durner, W. (2019). Numerical test of the laboratory evaporation method using coupled water, vapor and heat flow modelling. *Journal of Hydrology*, *570*(November 2018), 574–583. https://doi.org/10.1016/j.jhydrol.2018.12.045

Iden, S. C., & Durner, W. (2014). Comment on “Simple consistent models for water retention and hydraulic conductivity in the complete moisture range” by A. Peters. *Water Resources Research*, *50*(9), 7530–7534. https://doi.org/10.1002/2014WR015937

Peters, A. (2013). Simple consistent models for water retention and hydraulic conductivity in the complete moisture range. *Water Resources Research*, *49*(10), 6765–6780. https://doi.org/10.1002/wrcr.20548

Peters, A. (2014). Reply to comment by S. Iden and W. Durner on “Simple consistent models for water retention and hydraulic conductivity in the complete moisture range”. *Water Resources Research*, *50*(9), 7535–7539. https://doi.org/10.1002/2014WR016107

Peters, A., & Durner, W. (2008). Simplified evaporation method for determining soil hydraulic properties. *Journal of Hydrology*, *356*(1–2), 147–162. https://doi.org/10.1016/j.jhydrol.2008.04.016

Peters, A., Iden, S. C., & Durner, W. (2015). Revisiting the simplified evaporation method: Identification of hydraulic functions considering vapor, film and corner flow. *Journal of Hydrology*. https://doi.org/10.1016/j.jhydrol.2015.05.020

Schindler, U. (1980). Ein Schnellverfahren zur Messung der Wasser- leitfaähigkeit im teilgesä ttigten Boden an Stechzylinderproben. January, 1–7.

Šimůnek, J., M. Šejna, A., Saito, H., Sakai, M., & Genuchten, M. Th. V. (2013). The HYDRUS-1D Software Package for Simulating the Movement of Water, Heat, and Multiple Solutes in Variably Saturated Media. *Department of Environmental Sciences -*, *Version 4.*(June), 343–343.

3. CHAPTER III: A fitting procedure for the soil water retention function with improved uncertainty analysis

Abstract

Soil water retention and conductivity functions (RF and CF) need to be calibrated for numerically solving the Richards equation, commonly used in vadose-zone hydrological modeling. The determination of RF and CF parameters is normally performed using measured data of water content, suction, and hydraulic conductivity under equilibrium (static) or transient (dynamic) conditions. Parameter fitting is performed by nonlinear regression, resulting in average parameter values, their uncertainty (commonly expressed as standard error), and parameter correlations. The standard error can be evaluated from the fitting process. We evaluated the uncertainty of the fitted parameters and proposed transformations of the RF parameters with two objectives: (i) to reduce the uncertainty of the parameters by expressing the RF using water contents at two arbitrary suctions, instead of the traditional saturated (θ_s) and residual (θ_r) water content and (ii) to transform the remaining RF parameters to approach a normal distribution for the uncertainty of the parameter probability distribution (PPD). To describe the RF, we focused on the frequently used Van Genuchten model. The proposed transformations allow an improved understanding of the parameter probability distribution of the fitted parameters and a reduction of the parameter uncertainty. The results show that the transformations allowed the reduction of the uncertainty of parameters θ_s and θ_r, and were successful in obtaining a normal distribution for PPD in most of the analyzed scenarios, also increasing the convergence of fitting procedures which use derivatives to search for the parameter values which minimize the summed squared error. The statistical benefits acquired with the transformations are relatively low, however, they come at no extra cost and do not require additional experimental data, physical considerations, or model simplifications.

3.1. Introduction

The soil water retention function (RF) and hydraulic conductivity function (CF) are used in vadose zone hydrological models to numerically solve the Richards

equation. Most of the retention functions of these models have a common mathematical structure as

$$\theta(h) = \theta_r + (\theta_s - \theta_r) S(h) \quad [36]$$

where $S(h)$ is the effective saturation function, h [cm] is the suction (absolute value of the matric potential), θ_s [cm^3 cm^{-3}] is the saturated water content defined as the water content at h = 0 and θ_r [cm^3 cm^{-3}] is the residual water content, the water content at h = ∞.

A common model to describe the RF and CF is the Van Genuchten (1980) model. It links a mathematical function for RF with the capillary bundle model proposed by Mualem (1976) to yield a function for CF, and is referenced here as VGM. Other commonly used models are Van Genuchten (1980) coupled to the (Burdine, 1953) capillary bundle model, the Kosugi (1994, 1996) RF linked to the Kosugi (1999) CF, and the Groenevelt and Grant (Grant et al., 2010; Groenevelt & Grant, 2004) model linked to Mualem (1976) or Burdine (1953). All these models are composed of mathematical functions with mostly empirical fitting parameters.

For the VGM model, the $S(h)$ function is

$$S(h) = \left[1 + |\alpha h|^n\right]^{(1/n)-1} \quad [37]$$

and the CF is

$$K = K_s S(h)^l \left[1 - \left(1 - S(h)^{n/(n-1)}\right)^{1-(1/n)}\right]^2 \quad [38]$$

where α [cm^{-1}], n [–], and l [–], are fitting parameters and K_s [cm h^{-1}] is the saturated hydraulic conductivity. The VGM effective saturation model (Eq. 37) assumes unimodality, i.e., the retention function has exactly one inflection point. If the assumption of unimodality is satisfied, the reliable application of the VGM model depends on well-established parameters. Parameters K_s and θ_s have a clear physical meaning and may be acquired experimentally, although the determination may impose problems. For example, K_s presents a high spatial variability and extensive field experiments at the appropriate scale may be needed to measure its value properly (Durner, 1994; Pachepsky et al., 2014; Wen & Gómez-Hernández, 1996). Besides this,

measurements at full saturation may be difficult to establish (Durner, 1994). The θ_r parameter may be treated as a mere fitting parameter (Van Genuchten & Nielsen, 1985) or considered to have a physical meaning (Iden & Durner, 2014; Peters, 2013). These parameters, and the remaining ones, may be calibrated by fitting to experimental data of retention and/or (less common) hydraulic conductivity.

The most well-known software for specifically fitting the RF and CF to observed data is RETC (Van Genuchten et al., 1991). Alternatively, any statistical fitting software for generic models as well as dedicated libraries for different programming languages may be employed. The fitted parameters carry uncertainties that can be expressed as a probability density distribution, where the reported parameter value is the most likely one. A quick and frequently used method to obtain the best estimate for a parameter is to consider the uncertainty of the estimated parameter as a normal distribution. Consequently, minimizing the sum of squared deviations, named Least Squares method (LS), becomes equivalent to finding the Maximum Likelihood Estimator (MLE). In this condition, the uncertainties of the parameters can be expressed similarly to the standard deviation (σ) (Press, 2007). The parameter uncertainty will be affected by the dispersion of the fitting data, the ability of the model to reproduce the shape of the data set, and the statistical procedure used to perform the fitting.

The assumption of normality for the parameter probability distribution, PPD, for the fitted parameters might not be assured, which can lead to problems when using the respective estimate of σ, when σ is relatively big. Also, considering that the normal density distribution ranges from -∞ to +∞ with any real number in its domain, physically inconsistent values will be included in the domain for α (>0), n (>1), and K_s (>0). This inconsistency may become a problem when stochastically using the estimated uncertainty of the parameters to analyze the quality of the fitting procedure or evaluate a model compared to a data set, as in Wesseling et al. (2020a, 2020b) and Pinheiro & de Jong van Lier (2021).

Methods like Monte Carlo are commonly employed to generate the probability of the parameter value of a model, given a data set, yielding the complete probability distribution of the parameter. However, this technique requires much more computational effort. Expressing the parameter uncertainty as a single value (deviation) is usually preferred, but the normality of the frequency distribution of the

parameters should be warranted. Furthermore, it is of obvious interest to minimize parameter uncertainty.

In this context, we propose parameter transformations for the VGM parameters of the RF to improve the normality of the associated frequency distributions. We also propose a transformation of residual and saturated water contents allowing a reduction in the uncertainty of these parameters. The proposed parameter transformations allow an improved soil hydraulic fitting process, especially important in the setting of stochastic (Monte Carlo) modeling procedures.

3.2. Material and Methods

3.2.1. Transforming the retention function expression

To develop a new formulation of RF models based on Eq. 36, we consider

$$S(h) = \int_{-\infty}^{h} f(h) \qquad [39]$$

where $f(h)$, the base model equation (Eq. 36), can be expressed as

$$\theta(h) = \theta_r + (\theta_s - \theta_r)\left[\int_{-\infty}^{h_1} f(h) + \int_{h_1}^{h_2} f(h) + \int_{h_2}^{h} f(h)\right] \qquad [40]$$

This results in

$$\theta(h) = \theta_r + (\theta_s - \theta_r)\left[(S_1 - S_\infty) + (S_2 - S_1) + (S_h - S_2)\right] \qquad [41]$$

where S_1 is $S(h_1)$, S_2 is $S(h_2)$, S_h is $S(h)$ and S_∞ is $S(h)$ when h tends to infinity, corresponding to S_∞ tending to zero in the VGM model. Then

$$\theta(h) = \theta_r + (\theta_s - \theta_r)(S_1) + (\theta_s - \theta_r)(S_2 - S_1) + (\theta_s - \theta_r)(S_h - S_2) \qquad [42]$$

We now define two arbitrary water contents θ_1 and θ_2, corresponding to S_1 and S_2:

$$\theta_1 = \theta_r + (\theta_s - \theta_r)S_1 \qquad [43]$$

$$\theta_2 = \theta_r + (\theta_s - \theta_r)S_2 \qquad [44]$$

Isolating θ_r from Eq. 43 yields

$$\theta_r = \frac{\theta_1 - \theta_s S_1}{1 - S_1} \quad [45]$$

Substituting Eq. 45 in Eq. 44. and isolating θ_s results in

$$\theta_s = \frac{\theta_2 (1 - S_1) - \theta_1 (1 - S_2)}{S_2 - S_1} \quad [46]$$

and substituting Eq. 46 in Eq. 43 yields

$$\theta_r = \frac{\theta_1 S_2 - \theta_2 S_1}{S_2 - S_1} \quad [47]$$

Subtracting Eq. 46 from Eq. 45 allows to solve for (θ_s - θ_r) in terms of θ_2 and θ_1.

$$\theta_s - \theta_r = \frac{\theta_2 - \theta_1}{S_2 - S_1} \quad [48]$$

and substituting Eqs. 43 and 48 in Eq. 42 yields

$$\theta(h) = \theta_2 + \frac{\theta_2 - \theta_1}{S_2 - S_1}\left[S(h) - S_2\right] \quad [49]$$

Eq. 49 is analogous to Eq. 36, now expressing the RF in terms of any two arbitrary θ and h points, (θ_1, h_1) and (θ_2, h_2) instead of the original θ_r and θ_s. This formulation is an extension of the proposed RF equation presented by Inforsato et al. (2020). The conversion to (θ_s , 0) and (θ_1 , ∞) parameters can be performed with Eqs. 46 and 47.

3.2.2. Parameter transformations

The transformations used for the VGM parameters, originally proposed by Carsel & Parrish (1988) and applied by Peters and Durner, (2015), are

$$\begin{aligned} \kappa &= \ln(K_s) \Leftrightarrow K_s = e^{\kappa} \\ \tau &= \ln(\alpha) \Leftrightarrow \alpha = e^{\tau} \\ \nu &= \ln(n-1) \Leftrightarrow n = 1 + e^{\nu} \end{aligned} \quad [50]$$

Inserting Eq. 50 in Eq. 37 yields, for the RF

$$S(h)=\left[1+\left|e^{\tau}h\right|^{\left(1+e^{v}\right)}\right]^{\frac{-e^{v}}{1+e^{v}}} \quad [51]$$

For the CF we obtain, by substitution in Eq. 38

$$K=e^{\kappa}S(h)^{l}\left[1-\left(1-S(h)^{\frac{1+e^{v}}{e^{v}}}\right)^{\frac{e^{v}}{1+e^{v}}}\right]^{2} \quad [52]$$

It should be noticed that both τ and κ are affected by the dimension of α and K_s, respectively. For the case of τ, defining h_1 as suction expressed in unit ς_1 and h_2 as suction expressed in unit ς_2, corresponding values of α are α_1 (ς_1^{-1}) and α_2 (ς_2^{-1}). A conversion factor a (ς_2 ς_1^{-1}) can be determined such that,

$$h_2=ah_1 \Leftrightarrow a=\frac{h_2}{h_1}=\frac{\alpha_1}{\alpha_2}=\frac{e^{\tau_1}}{e^{\tau_2}} \quad [53]$$

from which it follows that

$$\tau_2=\tau_1-\ln(a) \quad [54]$$

For example, in a soil with α = 0.015 cm^{-1} = 1.5 m^{-1}, τ will be equal to ln(0.015) = -4.20 (for α in cm^{-1}) or to ln(1.5) = 0.405 (for α in m^{-1}). In this case, a = m^{-1}/cm^{-1} = 100, ln(a) = 4.605 and Eq. 38 is verified true.

Similarly, to convert between units ϕ_1 and ϕ_1 for K with conversion factor b (ϕ_2 ϕ_1^{-1}), for the conversion of κ we obtain

$$\kappa_2=\kappa_1-\ln(b) \quad [55]$$

To transform τ from its value τ_1 corresponding to unit ς_1 for h to τ_2 corresponding to unit ς_2, it can be shown that

$$\tau_2=\frac{\tau_1}{a} \quad [56]$$

$$\tau=\ln(\alpha) \Leftrightarrow \alpha=e^{\tau} \quad [57]$$

3.2.3. Software description

A software was developed to fit the proposed model to data using the Marquardt-Levenberg method, ML, (Marquardt 1963). The ML is widely used and well documented (Press, 2007). A brief explanation follows.

The ML fits the parameters of an established model by minimizing the sum of squares between the observed data and the predicted data through an iterative method, i.e. non-linear least squares. ML is a combination of the G*radient descent* method and *Newton's* method. In ML, each parameter step of the iteration to reduce the sum of squares is determined by

$$\mathbf{a}_{next} = \mathbf{a}_{current} + \left(\mathbf{J}^{\mathrm{T}} \cdot \mathbf{J} + \lambda \cdot \mathbf{I}\right)^{-1} \cdot \mathbf{J}^{\mathrm{T}} \cdot \mathbf{r} \qquad [58]$$

where $\mathbf{a}$ is the vector containing the parameter values, $\mathbf{J}$ is the Jacobian matrix, $\mathbf{I}$ is the identity matrix, $\mathbf{r}$ is the residues vector and λ is a damping factor. If λ is taken high, the iteration step tends to become equal to *Newton's* method, and if λ is assumed low, the iteration step tends to the *Gradient descent* method, considered more adequate for local adjustments. The superscripted letter T refers to the transposed matrix, the *current* subscript refers to current parameter values (fitted in the last iteration), and *next* refers to the values of the fitted parameters in the current iteration. In Eq. 58, the weight matrices are hidden, and the $\mathbf{J}^{\mathrm{T}} \cdot \mathbf{J}$ approximates the Hessian matrix, avoiding the need to use second-order derivatives.

After fitting, the parameter standard error, σ, is estimated by

$$\begin{aligned} \mathbf{C} &= \left(\mathbf{J}^{\mathrm{T}} \cdot \mathbf{J}\right)^{-1} \\ \sigma_j^2 &= \mathbf{C}_{jj} \end{aligned} \qquad [59]$$

where $\mathbf{C}$ is the variance-covariance matrix, the subscribed j refers to each component of the vector containing the standard error for each parameter and the jj refers to each component of the diagonal of the $\mathbf{C}$ matrix. Although this formula provides good predictions of the uncertainty when the deviations are low (when the estimation of the parameter uncertainty of the MLE is approximately equal to the parameter uncertainties provided by the LS method), the exactness tends to decrease with increasing deviation leading to an error in the uncertainty estimation, as the

determination of the parameter uncertainties predicted by Eq. 59 does not include biased parameter uncertainties.

3.2.4. Complex-step differentiation approximation

The derivatives used in the developed software are calculated using complex-step differentiation. The underlying theory can be found in Squire & Trapp (1998), Martins et al. (2003), and Anderson et al. (2001). Only a brief explanation based on Martins et al. (2003) is given here.

Considering that a complex function $f(z)=u(x,y)+i\cdot v(x,y)$ is differentiable in the imaginary plane, the Cauchy-Riemann equations are satisfied:

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y} \quad \& \quad \frac{\partial u}{\partial y}=-\frac{\partial v}{\partial x} \qquad [60]$$

The desired derivative is defined as

$$f'(z)=\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x} \qquad [61]$$

and, since the imaginary part is zero for this problem, the derivative becomes $f'(z)=\frac{\partial u}{\partial x}$. Using the leftmost expression from Eq. 60, the derivative can be defined as

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y}\rightarrow f'(z)=\frac{\partial v}{\partial y}\rightarrow f'(z)=\lim_{h^*\to 0}\frac{v(x,y+h^*)-v(x,y)}{h^*}=\lim_{h^*\to 0}\frac{\mathrm{Im}\left[f(x+ih^*)\right]}{h^*} \qquad [62]$$

where Im refers to the real value multiplied by i of the complex number. Then, the rightmost part of Eq. [62] can be approached by

$$f'(x)\approx\frac{\mathrm{Im}\left[f(x+ih^*)\right]}{h^*} \quad \left(\text{for small values of } h^*\right) \qquad [63]$$

Eq. 63 is very convenient for our purposes as it allows to be applied to other RF and CF models. It is computationally more efficient than calculating the analytical derivatives, and it does not require to calculate the differences used for finite

differences numerical differentiation $f(x+h)-f(x)$ or $f(x+h)-f(x-h)$, which may represent a source of error.

3.2.5. Parameter uncertainty analysis and comparison

- Developed scenarios: criteria and analysis

To evaluate the fitting procedure, synthetic data were used to obtain better control of the fits and to exclude errors arising from the relationship between the model and the fitted data. Three parameter sets for RF available in RETC software (Van Genuchten et al., 1991) were used for analysis, corresponding to Clay Loam (CL), Sandy Loam (SaL), and Silty Clay (SiC) USDA texture classes (Table 6). For each of the three RF parameters, water contents were calculated at suctions selected equidistantly on the pF-scale (pF = $\log_{10}|h[\text{cm}]|$), step 0.3, from pF 0.0 to pF 4.2, totaling 15 (θ, h) data pairs.

Table 6. Parameters of the retention functions analyzed for the VGM model.

	θ_r	θ_s	α [cm^{-1}]	n	K_s [cm h^{-1}]	λ
Clay Loam	0.095	0.41	0.019	1.31	0.26000	0.5
Sandy Loam	0.065	0.41	0.075	1.89	4.42083	0.5
Silty Clay	0.070	0.36	0.005	1.09	0.02000	0.5

Noise was added to the calculated water content data to allow for the evaluation of the parameter uncertainty resulting from the fitting procedure. Scenarios with different noise levels were applied, corresponding to the calculated water contents at the respective suction together with four values of standard deviation σ. For each scenario, the original water contents in each dataset were substituted by a biased one (θ_{bias}), maintaining the respective h value. The θ_{bias} was randomly sampled, obeying a normal distribution with respective θ as mean and deviation corresponding to the four considered deviation levels, σ, equal to 0.001, 0.005, 0.01, or 0.02 cm^3 cm^{-3}, respectively).

- Bootstrap method

The Resampling Bootstrap Method (Efron & Tibshirani, 1993; Press, 2007) was used to create the probability distribution of parameters. The method consists of substituting the dataset with randomly chosen data from the same dataset, maintaining its original size. This step was repeated 10,000 times for each analyzed scenario. After the Bootstrap process, each of the sampled data sets was used to fit the RF, resulting in 10,000 new sets of fitted parameters used to create the probability distributions. The analysis of transformed parameters was performed for transformed parameters α and n (with Eq. 49), and using all the transformations (α, n, as well as substituting θ_s and θ_r by two arbitrarily "anchored" points, with Eq. 50).

- Normality test

To determine the normality of the parameter probability distribution (PPD), the D'Agostino normality test (D'Agostino, 1971; D'Agostino & Pearson, 1973) was applied. D'Agostino's normality test (D'Ag.) considers skewness (Sk.) and kurtosis (Kur.) and as such it compares the bias and the tail of the distribution to a normal distribution. The skewness and kurtosis were also calculated for all PPD, subtracting 3 from the kurtosis, making it zero for a normal distribution.

To visually analyze the difference between the predicted normal distributions and the PPD provided by the Bootstrap Method, histograms with PPD were plotted (50 bins in each graph, excluding data outside the interval +/- 3 σ), as well as each respective normal distribution (deviation and average provided by the fitting procedure).

For statistical validation of the observed differences, the Wilcoxon Signed Ranks test was used (Conover, 1999).

3.3. Results and Discussion

3.3.1. Description of the generated soils

Figure 10 presents the exact retention curves for the three soils used as bases (Table 6), together with the generated data points for each scenario (σ equal to 0.001, 0.005, 0.01 and 0.02 $cm^3\ cm^{-3}$).

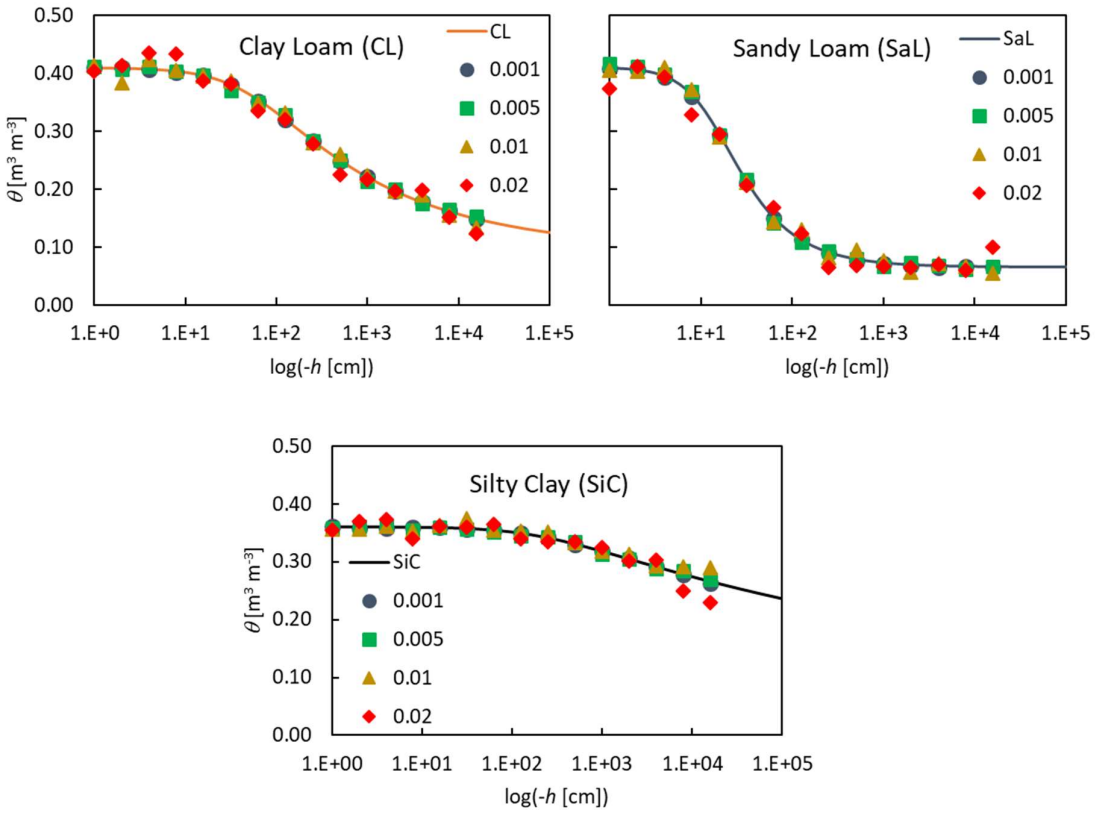

Figure 10. Plotted functions of the three analyzed soils: Clay Loam (CL), Sandy Loam (SaL) and Silty Clay (SiC). The solid line represents the exact VGM retention function, the dots represent the generated data at four levels of added noise.

3.3.2. Arbitrary anchoring point analysis

The use of Eq. 49 instead of the traditional Eq. 36 allows to establish one or two well-determined RF points, (θ, h), and to use them as know parameters, reducing the number of parameters to be fitted. Fig 11 shows the constraints of the VGM model for some values of parameters n and α, using Eq. 49 and with parameters (θ_2, h_2) = (0.30, 10^2) and (θ_1, h_1) = (0.15, 10^4).

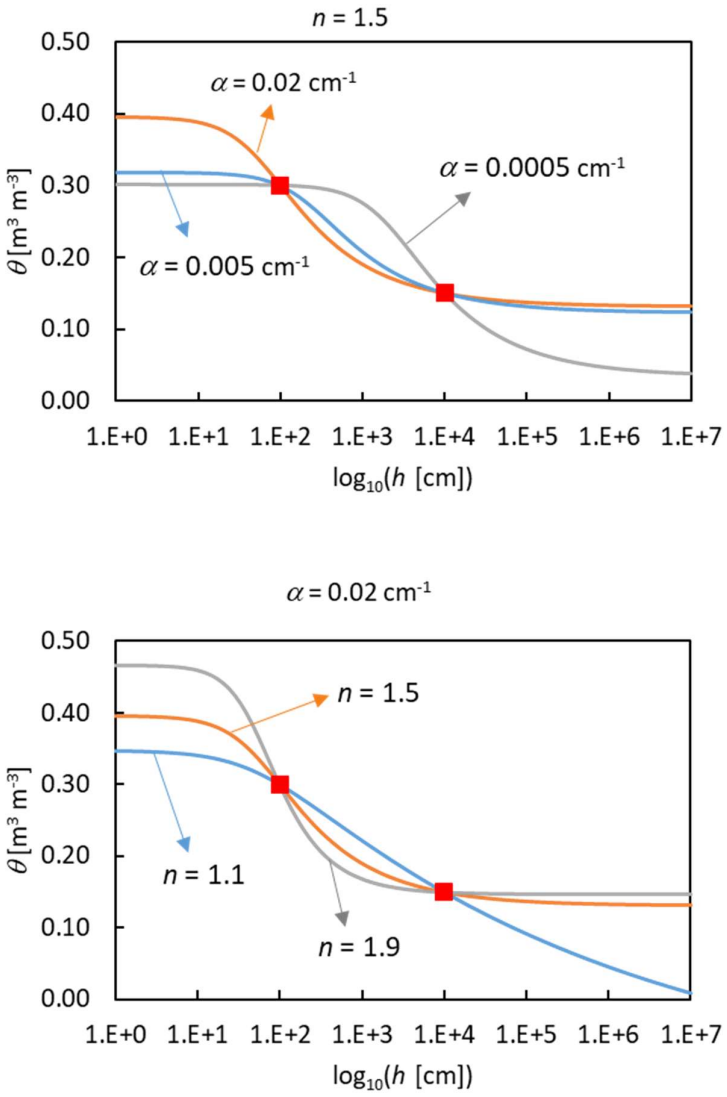

Figure 11. Soil water retention curves expressed by Eq. 49 with parameters (θ_2, h_2) = (0.30, 10^2) and (θ_1, h_1) = (0.15, 10^4), varying α at n = 1.5 (top), and varying n at α = 0.02 cm^{-1} (bottom).

3.3.3. Anchoring point and parameter uncertainty

Table 7 shows traditionally fitted parameters θ_s and θ_r for the Silty Clay soil at the level of 0.01 cm^3 cm^{-3} added noise (SiC 0.01), together with results for fits using Eq. 49 with alternative anchoring points. In transform 1, h_1 was set to 1 cm (pF = 0) and h_2 was set to 15 849 cm (pF = 4.2), whereas in transform 2 these values were

h_1 = 4 cm (pF = 0.6) and h_2 = 7 943 cm (pF = 3.9). The results show the deviation, σ, of the parameter estimates to decrease, especially at the dry side (θ_2 versus θ_r) when using anchoring points within the observed suction domain.

Table 7. Comparison of the deviation, σ, of the model parameters θ_s and θ_r, and parameters of the model expressed in two arbitrary points (θ_2, h_2) and (θ_1, h_1) for the fitting procedure made with data from the Silty Clay at the level of 0.01 cm^3 cm^{-3} added noise (SiC 0.01), h values in cm.

	Traditional		Transform 1		Transform 2	
	h_2= 0.0	h_1=∞	h_2=1.0	h_1=15849	h_2=4.0	h_1=7943
	θ_s	θ_r	θ_2	θ_1	θ_2	θ_1
value	0.360	0.282	0.360	0.288	0.360	0.291
σ	$2.33 \cdot 10^{-3}$	$1.06 \cdot 10^{-2}$	$2.33 \cdot 10^{-3}$	$4.71 \cdot 10^{-3}$	$2.32 \cdot 10^{-3}$	$3.40 \cdot 10^{-3}$

To illustrate and understand the deviations observed in Table 7, for each soil with added noise we performed 1000 regressions, each one using a different value of h_1 according to an even distribution on a log-scale between 1 cm and 10^7 cm. On the left, for each h_1 value, a corresponding θ_1 was fitted, maintaining the value of h_2 fixed at 1.0. The calculated deviation of θ_1 (σ) is shown in Figure 12 for the cases of Clay Loam 0.02, Sandy loam 0.02 and Silty Clay 0.01. In the case of Clay Loam 0.02, the magnitude of the deviation is considerably lower when choosing h_1 inside the measured data domain and $\sigma(\theta_1)$ increases steadily when $|h_1|$ is chosen at values higher than 15 000 cm. Similarly, on the right part of the Figure 12 the deviation (σ) of θ_2 as a function of h_2 while fixing h_1 at 10^7 is shown. An h_2 within the observed range of suctions reduces σ, while the choice of an h_2 at a large (or “infinite”) value, like in the case of traditional θ_r, yields a high parameter uncertainty.

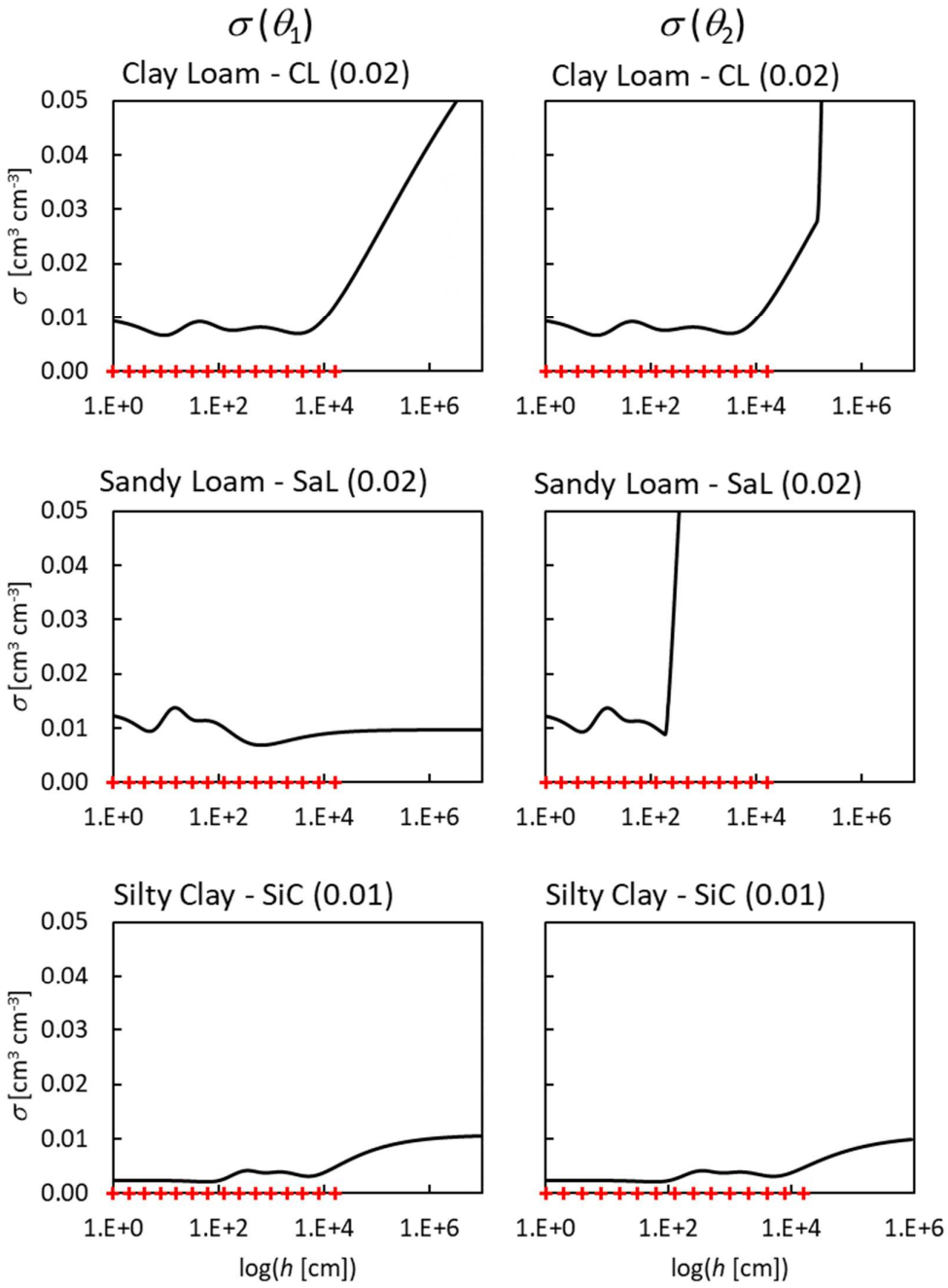

Figure 12. Left side: deviations (σ) of fitted θ_1 as a function of h_1 for scenarios Clay Loam 0.02, Sandy Loam 0.02, and Silty Clay 0.01, maintaining h_2 fixed at 1.0. Right side: deviations (σ) of fitted θ_2 for each value h_2 for the same soil data, maintaining h_1 fixed at 10^7 cm. Red cross-marks on the h-axis indicate the observed suctions. The crosses represent the observed data for each soil scenario, Clay Loam 0.02, Sandy Loam 0.02, and Silty Clay 0.01. The black continuous lines are the deviations at each specific h.

Figure 13 gives another example of the described behavior, showing the observed data and the fitted RF. In this figure, the dashed lines represent the upper and lower deviations (σ), estimated at each value of h with a new regression considering the arbitrary h_1 value as the respective suction.

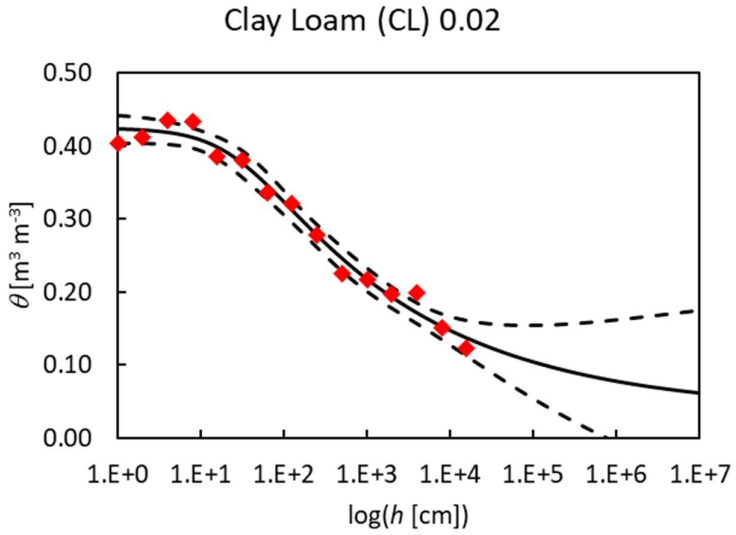

Figure 13. Retention function with the deviation of the parameter at each value of h. Red diamonds represent the observed data (θ-h), the continuous black line represents the fitted retention function, dashed lines are the upper and lower deviations (σ) calculated at each h value (fixing each h as h_1).

The best choice for h_1 and h_2 in Eq. 49 to minimize respective values of σ corresponds to low values of the second derivative of the objective function with respect to θ_2 or θ_1, since ML uses an approximation of second derivatives to estimate σ. This will lead to different optimal values of h_2 and h_1 depending on the data set.

There is no straightforward procedure to determine the best values of h_2 and h_1 to minimize σ, and a trial-and-error procedure is the best option. However, from our tests, it follows that the values for h_2 and h_1 can best be chosen inside the range of observed data, slightly distant from the highest and lowest observed h boundaries.

3.3.4. Parameter normality test analysis

- Transforms of parameters α and n

Parameter transformations according to Eq. 50 include RF parameters $\tau = \ln(\alpha)$ and $\nu = \ln(n\text{-}1)$. Applying D'Agostino's normality test, skewness and kurtosis were determined for the transformed parameters τ and ν distribution obtained by bootstrap for each of the soil data scenarios (Table 8). A low D'Agostino test value indicates proximity to a normal density distribution. In general, the parameter transformations approached the distributions to normal ones, with some exceptions (Table 8). Table 9 presents the same results for untransformed parameters. The differences between the D'Agostino's test for transformed n and α are significant according to the Wilcoxon test (Wilcoxon = 28.0, p-value = $6.94 \cdot 10^{-4}$), for the skewness (Wilcoxon = 8.0, p-value = $1.19 \cdot 10^{-5}$) and kurtosis (Wilcoxon = 62.0, p-value = 0.036). A general transformation that improves every PPD was not found, nevertheless, the results show improvement in approaching the PPD to the normal distribution. When the transform moves the PPD away from the normal distribution, the absolute difference value of D'Agostino's test tends to be lower when compared with the scenarios where the transformation approached the PPD to normal distribution.

The p-value for D'Agostino's test was calculated for skewness and kurtosis, again to compare the transformed and untransformed parameters, but none of those led to a significant value. Skewness values of transformed parameters, τ, and ν, are in between +/- 0.5, and the kurtosis in between +/- 1.5. A positive kurtosis indicates that considering σ as the deviation will overestimate the true deviation.

Table 8. D'Agostino's (D'ag.) normality test values, skewness (Sk.) and kurtosis (Kur.) of soil for the distribution of the uncertainty of the parameters τ and ν. Highlighted values refer to cases were normality decreased when applying the proposed transformations of α and n; non-highlighted values refer to an approximation of the normal distribution when applying the transformations.

		Parameter					
		$\tau = \ln(\alpha\,[\mathrm{cm}^{-1}])$			$\nu = \ln(n\text{-}1)$		
Soil (σ)		D'ag.	Sk.	Kur.	D'ag.	Sk.	Kur.
CL	(0.001)	86.6	-0.12	-0.32	5.2	0.03	-0.10
CL	(0.005)	117.9	0.27	-0.08	99.7	0.03	0.63
CL	(0.01)	117.9	0.27	-0.08	99.7	0.03	0.63
CL	(0.02)	819.5	0.69	0.99	61.6	-0.16	0.24
SaL	(0.001)	88.4	0.23	0.15	16.2	0.09	0.11
SaL	(0.005)	407.5	-0.41	0.91	178.3	-0.34	-0.11
SaL	(0.01)	173.1	0.32	0.28	107.8	-0.17	-0.33
SaL	(0.02)	127.4	0.26	0.28	132.9	0.26	0.30
SiC	(0.001)	260.8	0.41	0.01	13.4	-0.09	-0.01
SiC	(0.005)	663.6	0.25	-0.72	266.8	-0.45	0.17
SiC	(0.01)	357.1	0.30	1.12	88.0	0.13	0.46

Table 9. D'Agostino's (D'ag.) normality test values, skewness (Sk.) and kurtosis (Kur.) of soil for the distribution of the uncertainty of the parameters α and n.

		Parameter					
		$\alpha\,[\mathrm{cm}^{-1}]$			n		
Soil (σ)		D'ag.	Sk.	Kur.	D'ag.	Sk.	Kur.
CL	(0.001)	67.0	-0.09	-0.31	13.1	0.07	-0.11
CL	(0.005)	750.7	0.76	0.18	732.7	0.67	0.75
CL	(0.01)	750.7	0.76	0.18	732.7	0.67	0.75
CL	(0.02)	1896.5	1.26	1.66	438.7	0.56	0.12
SaL	(0.001)	102.3	0.24	0.17	21.2	0.10	0.10
SaL	(0.005)	356.6	-0.38	0.82	149.5	-0.29	-0.17
SaL	(0.01)	342.1	0.46	0.37	57.2	-0.01	-0.31
SaL	(0.02)	167.8	0.26	0.50	488.6	0.58	0.34
SiC	(0.001)	495.2	0.59	0.13	261.0	0.42	-0.19
SiC	(0.005)	678.6	0.68	-0.36	477.8	0.49	-0.45
SiC	(0.01)	497.9	0.23	1.82	703.1	0.66	0.76

Figures 14a and 14b present some distributions to visualize how the parameter distribution behaves and how the proposed transformations affect the uncertainty

distribution. Considering all fitted α parameters and $\tau = \ln(\alpha)$ parameters, the α parameter acquired with CL (0.02) data was the most distant from a normal distribution according to D'Agostino's test = 1896. The corresponding transformation approached the uncertainty to a normal distribution, significantly decreasing the absolute value of skewness and kurtosis. For the n parameter, its value acquired with CL (0.005) data was the most distant from a normal distribution according to D'Agostino's test, and the skewness reduced from 0.67 to 0.03 due to the transformation of α. Considering all four SaL datasets with added noise, SaL (0.005) was most distant from a normal distribution, for τ (transformed α), D'Agostino = 407, and SaL (0.02) for the parameter *n*, with D'Agostino = 489. Considering α and τ parameters for the SiC soil data with noise, the PPD most distant from a normal distribution was SiC (0.005), and considering *n* and $\nu = \ln(n\text{-}1)$, SiC (0.01). Unlike for the transformed parameter, an increase of probability near the minimum possible value for parameter *n* (where *n* = 1) in SiC (0.01) can be observed.

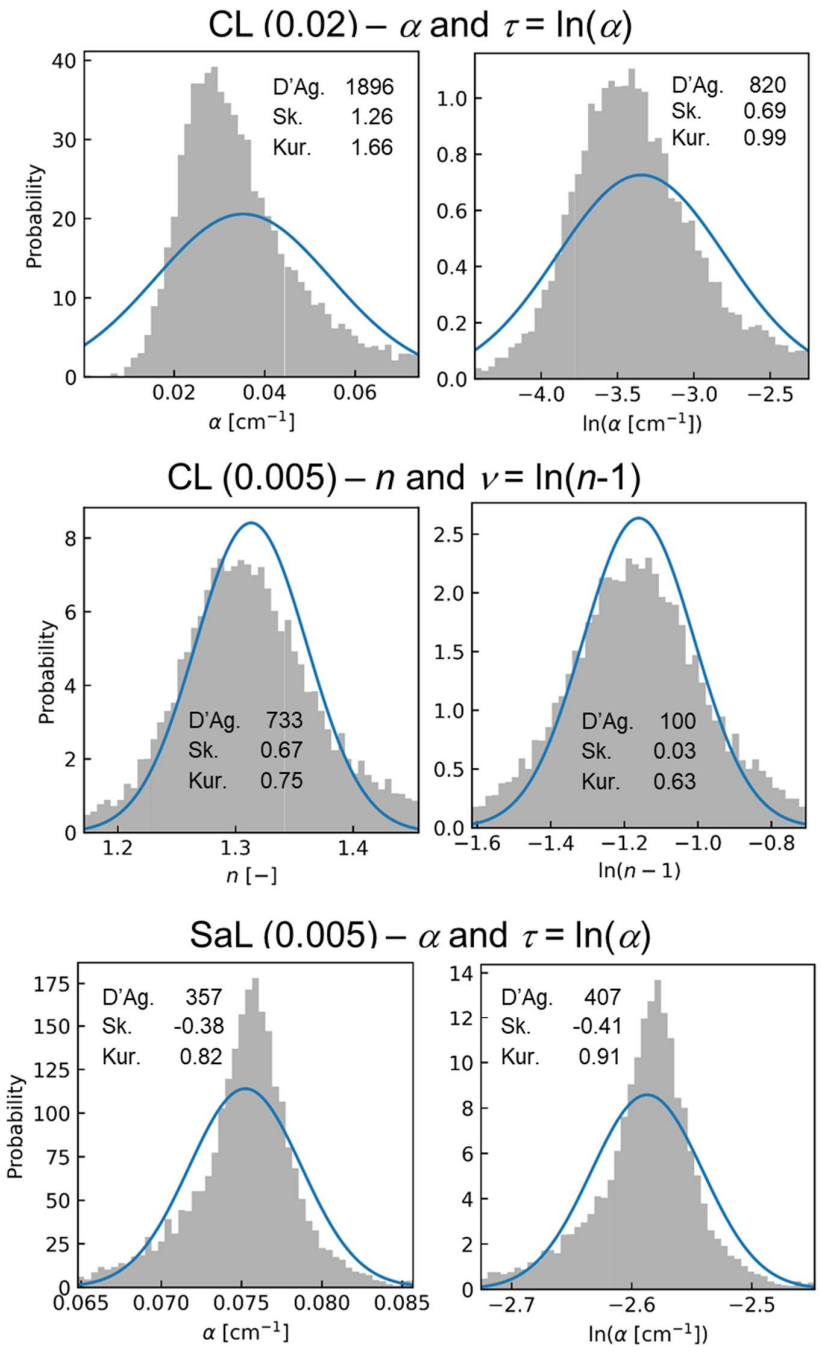

Figure 14a. Parameter probability per parameter value. The blue line represents the distributions of the uncertainty of the parameters considering calculated deviation σ. The normalized histogram is the parameter probability distribution obtained through bootstrap. D'Ag. Sk. and Kur. refer to D'Agostino's test, skewness and kurtosis of the gray density distribution.

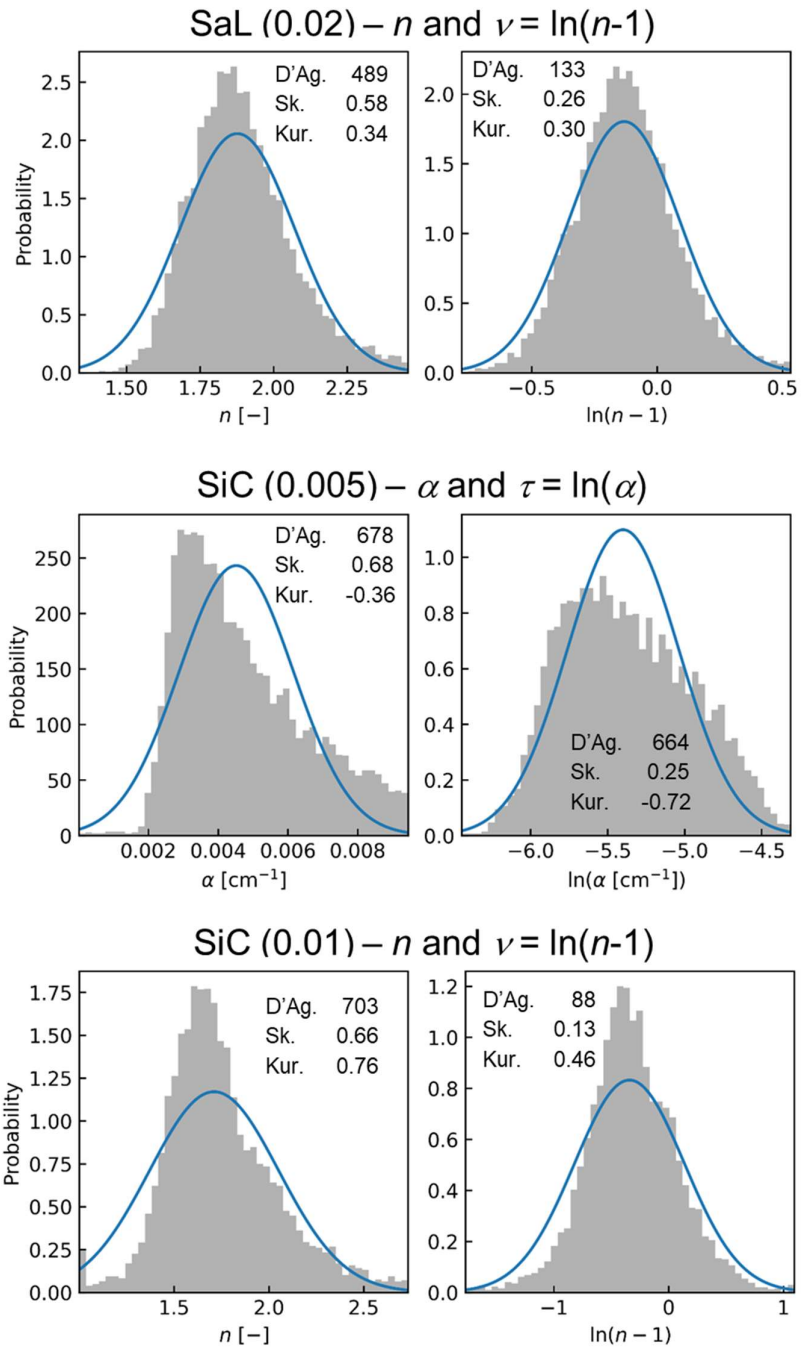

Figure 14b. Parameter probability per parameter value. The blue line represents the distributions of the uncertainty of the parameters considering calculated deviation σ. The normalized histogram is the parameter probability distribution obtained through bootstrap. D'Ag. Sk. and Kur. refer to D'Agostino's test, skewness and kurtosis of the gray density distribution.

- Retention function fitting analysis with simultaneous transform of all parameters

The effect of retention function fitting of a complete transform of all parameters ([α, n, θ_s and θ_r] to [τ, ν, θ_2 and θ_1]) was analyzed. To do so, two values for h_2 and h_1 were chosen following the suggestions described in the topic "*Anchoring point and parameter uncertainty*". Although optimum values for h could be found by analyzing Figure 12, this kind of information is not expected to be available before commonly performed fitting procedures. Therefore, values in between the fourth and fifth highest and lowest observed data were chosen, corresponding to h_2 = 11.2 cm (pF = 1.05) and h_2 = 1413 cm (pF = 3.15), applied to all the soils. Table 10 provides the D'Agostino normality test values of the PPD for all the parameters transformed simultaneously, indicating if the normality of the distribution increased (uncolored) or decreased (shaded). On average, the results show a higher degree of normality when transforming the parameters. Similarly, Table 11 presents the D'Agostino normality test values for parameter θ_s and θ_r for comparison with transformed values in Table 10 (D'Agostino´s test for values of n and α is shown in Table 9).

Table 10. D'Agostino's normality test values for the probability distribution of the transformed parameters (all transformations performed simultaneously), for the fitting procedure of analyzed soil data. A shaded background indicates a lower degree of normality after parameter transformation, and unshaded cells represent indicate a higher degree of normality after transformation. The value for h_2 was pF = 1.05 and for h_1 pF = 3.15.

	D'Agostino			
Soil (σ)	$\tau = \ln(\alpha\ [cm^{-1}])$	$\nu = \ln(n\text{-}1)$	θ_2	θ_1
CL (0.001)	83.8	4.8	161.2	128.7
CL (0.005)	316.0	34.0	45.2	26.8
CL (0.01)	129.9	136.6	186.4	439.7
CL (0.02)	698.8	61.6	19.4	16.7
SaL (0.001)	72.1	27.4	165.7	23.9
SaL (0.005)	531.6	155.3	934.4	18.9
SaL (0.01)	205.5	107.4	145.3	62.7
SaL (0.02)	161.4	99.8	3.7	42.1
SiC (0.001)	328.5	27.8	33.9	22.2
SiC (0.005)	595.4	241.7	384.5	66.4
SiC (0.01)	329.0	142.4	223.8	304.8

Table 11. D'Agostino's normality test values for the probability distribution of parameters θ_s and θ_r, for the fitting procedure of analyzed soil data.

	D'Agostino	
Soil (σ)	θ_s	θ_r
CL (0.001)	446.1	25.8
CL (0.005)	337.6	592.2
CL (0.01)	337.6	592.2
CL (0.02)	461.7	635.6
SaL (0.001)	616.9	121.6
SaL (0.005)	1259.7	13.4
SaL (0.01)	2204.1	78.5
SaL (0.02)	114.6	100.3
SiC (0.001)	68.2	892.6
SiC (0.005)	201.3	1584.0
SiC (0.01)	182.4	789.5

The values of the D'Agostino tests for normality confirm, on average, an increase in the degree of normality of the distribution of the PPD of the parameters in the performance of the fitting. The difference has statistical significance according to the Wilcoxon test (Wilcoxon = 94.0 and *p*-value = $2.87 \cdot 10^{-6}$). Some differences between the values of the D'Agostino normality test are observed when fitting transformed α and n only (Table 8) and performing the fitting procedure transforming all parameters (Table 10), which indicates that transforming the θ_s and θ_r parameters impacts the normality of uncertainty of other parameters. Nonetheless, this change is not significant (Wilcoxon = 116.0 with *p*-value = 0.75).

Figures 15a, 15b, and 15c present the results of the fitting process of the abovementioned scenarios, visually demonstrating the impact of transformations $\theta_s \leftrightarrow \theta_2$ and $\theta_r \leftrightarrow \theta_1$. The soils presented in these figures are CL (0.02), SaL (0.02), SiC (0.005) for parameters θ_s and θ_2; and CL (0.001), SaL (0.01), SaL (0.005) and SiC (0.01) for parameters θ_r and θ_1. Transformed and untransformed parameters α and *n* are not shown due to visual similarities to Figures 14a and 14b. The blue lines correspond to the normal distributions, i.e., the estimated PPD for each parameter estimated with the LS fitting procedure. The gray histograms are the PPD provided by the bootstrapping method. Close to each PPD, the D'ag. Sk. and Kur. values refer to D'Agostino's test, skewness, and kurtosis of the respective histogram. Above each pair

of histograms, the acronym and level of noise added for each soil dataset are presented. The chosen soils synthesize the visual characteristics observed when comparing transformed and untransformed parameters. The $\theta_s \leftrightarrow \theta_2$ transformation PPD approximated to a normal distribution in soils dataset CL (0.02) and SaL (0.02), and was distanced in SiC (0.005). With the $\theta_r \leftrightarrow \theta_1$ transformation, the PPD approximated in soils datasets SaL (0.01) and SiC (0.01), but was distanced in soils CL (0.001) and SaL (0.005). In CL (0.02) and SaL (0.02) a visual approximation to a normal distribution is observed, which agrees with the result of the D'Agostino test. In SiC (0.005) the distance is increased mainly because of the kurtosis. Soil CL (0.001) was distanced from a normal distribution by transforming the parameter. For SaL (0.01) and SaL (0.005) only a small difference is observed, the first one coming closer and the second one being distanced from the normal distribution respectively. The SiC (0.01) showed more normality concerning skewness and kurtosis, but the transformed PPD had multimodal behavior, which may be a consequence of the Bootstrapping and a low amount of data in the in the drier part of the function. Although in some soils presented in Figures 15a, b, and c the transformation decreased the normality of the PPD, in all these scenarios the transformation narrowed the distribution.

Considering all the transformed parameters, parameter θ_1 for soil CL (0.02) showed the closest-to-zero skewness (Sk. = 0.003), and SaL (0.005) for parameter θ_2 the highest one (Sk. = 0.77), with an average skewness of 0.049. The lowest kurtosis (Kur. = 0.005) occurred for SiC (0.001) and parameter $\nu = \ln(n\text{-}1)$, and the highest one (Kur. = 1.46) for soil CL (0.01) for the same parameter, with an average kurtosis of 0.17. For comparison, the average skewness and kurtosis for the untransformed parameters are 0.12 and 0.38 respectively. Most of the values obtained for skewness and kurtosis for transformed parameters may be considered close to a normal distribution.

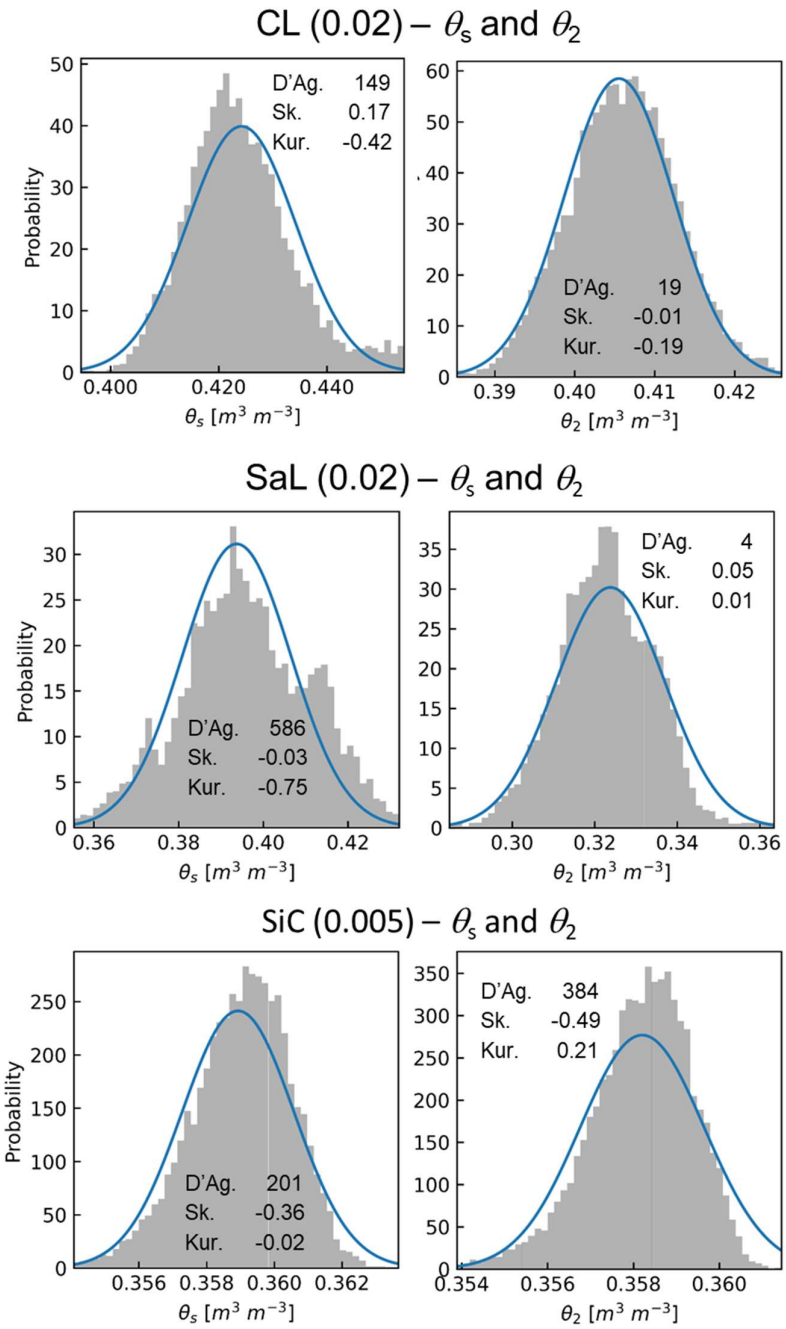

Figure 15a. Frequency distribution of the parameters θ_s and θ_2, estimated from the fitting procedures for soil scenarios CL(0.02), SaL(0.02), and SiC (0.005). The normalized histogram is the parameter probability distribution obtained through Bootstrapping, the blue curve represents the corresponding normal distribution. D'ag. Sk. and Kur. refer to D'Agostino's test, skewness, and kurtosis of the gray density distribution.

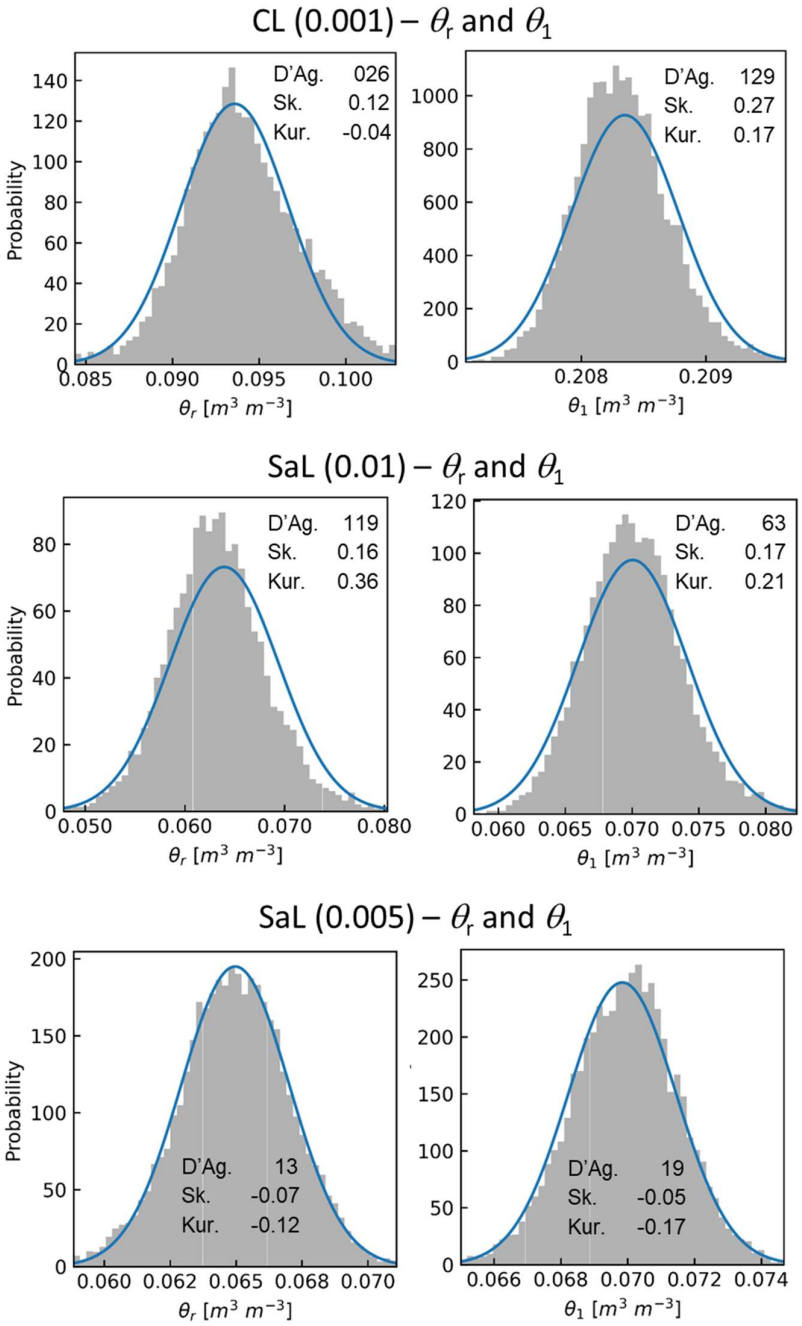

Figure 15b. Frequency distribution of the parameters θ_r and θ_1, estimated from the fitting procedures for soil scenarios CL(0.001), SaL(0.01), and SaL (0.005). The normalized histogram is the parameter probability distribution obtained through Bootstrapping, the blue curve represents the corresponding normal distribution. D'ag. Sk. and Kur. refer to D'Agostino's test, skewness, and kurtosis of the gray density distribution.

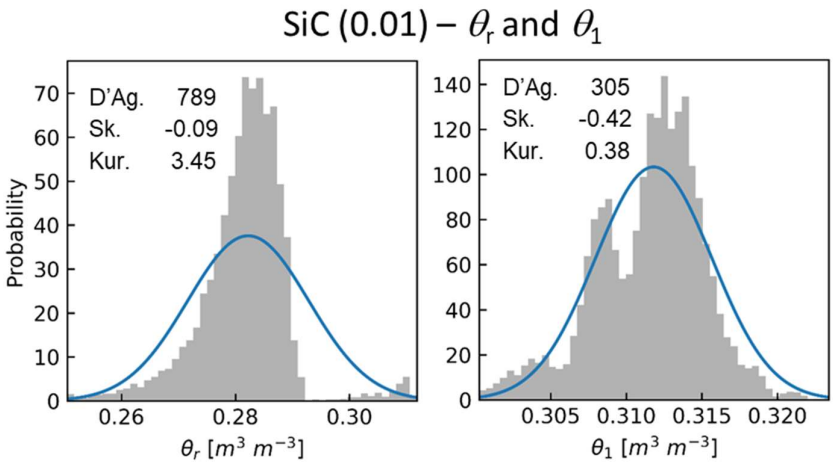

Figure 15c. Frequency distribution of the parameters θ_r and θ_1, estimated from the fitting procedures for soil scenarios SiC (0.01). The normalized histogram is the parameter probability distribution obtained through Bootstrapping, the blue curve represents the corresponding normal distribution. D'ag. Sk. and Kur. refer to D'Agostino's test, skewness, and kurtosis of the gray density distribution.

To test if the benefits of parameter transformation are significant only on soil data with higher statistical errors, a similar test was conducted for soil data with added noise levels 0.001 and 0.005 with untransformed parameters, and with a simultaneous transform of all parameters. The comparison between PPD with all parameters transformed and untransformed resulted in Wilcoxon = 46.0 with *p*-value = $2.0 \cdot 10^{-3}$, resulting in a statistically significant difference closer to the normal distribution when transformed (on average).

3.3.5. Correlation analysis

The correlation between the parameters is affected by the transformations. An analysis of the correlation was conducted for the fitting process considering untransformed and transformed parameters, ([α, n, θ_s and θ_r] or [τ , ν, θ_2 and θ_1] respectively) for all considered soil data sets. In general, a decrease in the absolute values of the correlation matrix (i.e., a lower correlation between parameters) is observed. Low linear correlation is desired because it increases the level of significance of the parameters, on the other hand, correlation close to 1 positive or negative, are indicators that the data can be explained without the need for one of the

parameters. Only the (α , n) parameter pair with its correlated transformation (τ , ν) were not statistically different in terms of correlation.

Wilcoxon tests were realized for correlations of all parameter pairs to determine the significance of difference between transformed and untransformed parameters. The results of the test reached the same value of *p*-value = $9.77 \cdot 10^{-4}$, showing a significant and systematic decrease in the correlations. Considering only the absolute values of the correlations, the parameter pairs presenting significant difference from their transformed equivalents are (θ_s , θ_r), (α , θ_r), (n , θ_r), maintaining the *p*-value = $9.77 \cdot 10^{-4}$. Table 12 provides the average of the absolute value of the pair correlations.

Table 12. Pairs of parameters with each respective average of the absolute value of their linear correlations ($corr$).

Untransformed		Transformed	
Parameter pair	$\overline{\lvert corr \rvert}$	Parameter pair	$\overline{\lvert corr \rvert}$
(α , n)	0.814	(τ , ν)	0.814
(α , θ_s)	0.645	(τ , θ_2)	0.454
(α , θ_r)	0.567	(τ , θ_1)	0.238
(n , θ_s)	0.419	(ν , θ_2)	0.390
(n , θ_r)	0.848	(ν , θ_1)	0.509
(θ_s , θ_r)	0.303	(θ_2 , θ_1)	0.084

A software to perform the nonlinear regressions with transformed and untransformed parameters as described in this chapter is available at https://github.com/infoleon/hp_Fit.

3.4. Conclusion

This study proposes transformations of the parameters of VGM retention function model with the purpose of obtaining parameter probability distributions closer to normal distributions and to decrease the uncertainty of the probability distribution of parameters θ_s and θ_r. With some considerations, the transformation for parameters θ_s and θ_r is applicable to any S-shaped retention function. It can be concluded that:

1. The transformations approach the parameter probability distributions to a normal distribution in most of the tested soil scenarios. Skewness and kurtosis of transformed parameters indicate normality of the distributions.
2. The proposed parameter transformations increase the convergence of the parameter in the fitting procedure when using the Marquardt method.
3. The traditionally used values of θ_s and θ_r in common retention functions can be substituted by two water contents at arbitrary pressure heads which, if chosen within the observed range of pressure heads, allow a reduction of parameter uncertainty.

3.5. References (Chapter III)

Anderson, W. K., Newman, J. C., Whitfield, D. L., & Nielsen, E. J. (2001). Sensitivity Analysis for Navier-Stokes Equations on Unstructured Meshes Using Complex Variables. AIAA Journal, 39(1), 56–63. https://doi.org/10.2514/2.1270

Burdine, N. T. (1953). Relative permeability calculations from pore size distribution data. Journal of Petroleum Technology, 5(3), 71–78. https://doi.org/10.2118/225-G

Carsel, R. F., & Parrish, R. S. (1988). Developing joint probability distributions of soil water retention characteristics. Water Resources Research, 24(5), 755–769. https://doi.org/10.1029/WR024i005p00755

Conover, W. J. (1999). Practical Nonparametric Statistics (3o ed). John Wiley & Sons, Inc. https://www.wiley.com/en-us/Practical+Nonparametric+Statistics%2C+3rd+Edition-p-9780471160687

D'Agostino, R. B. (1971). An omnibus test of normality for moderate and large size samples. Biometrika, 58(2), 341–348. https://doi.org/10.1093/biomet/58.2.341

D'Agostino, R., & Pearson, E. S. (1973). Tests for departure from normality. Empirical results for the distributions of b 2 and √ b 1. Biometrika, 60(3), 613–622. https://doi.org/10.1093/biomet/60.3.613

Durner, W. (1994). Hydraulic conductivity estimation for soils with heterogeneous pore structure. Water Resources Research, 30(2), 211–223. https://doi.org/10.1029/93WR02676

Efron, B., & Tibshirani, R. (1993). An introduction to the bootstrap. Chapman & Hall.

Grant, C. D., Groenevelt, P. H., & Robinson, N. I. (2010). Application of the Groenevelt-Grant soil water retention model to predict the hydraulic conductivity. Australian Journal of Soil Research, 48(5), 447–458. https://doi.org/10.1071/SR09198

Groenevelt, P. H., & Grant, C. D. (2004). A new model for the soil-water retention curve that solves the problem of residual water contents. European Journal of Soil Science, 55(3), 479–485. https://doi.org/10.1111/j.1365-2389.2004.00617.x

Iden, S. C., & Durner, W. (2014). Comment on “Simple consistent models for water retention and hydraulic conductivity in the complete moisture range” by A. Peters. Water Resources Research, 50(9), 7530–7534. https://doi.org/10.1002/2014WR015937

Inforsato, L., de Jong van Lier, Q., & Pinheiro, E. A. R. (2020). An extension of water retention and conductivity functions to dryness. Soil Science Society of America Journal, 84(1), 45–52. https://doi.org/10.1002/saj2.20014

Kosugi, K. (1994). Three-parameter lognormal distribution model for soil water retention. Water Resources Research. https://doi.org/10.1029/93WR02931

Kosugi, K. (1996). Lognormal Distribution Model for Unsaturated Soil Hydraulic Properties. Water Resources Research, 32(9), 2697–2703. https://doi.org/10.1029/96WR01776

Kosugi, K. (1999). General Model for Unsaturated Hydraulic Conductivity for Soils with Lognormal Pore-Size Distribution. Soil Science Society of America Journal, 63(2), 270–270. https://doi.org/10.2136/sssaj1999.03615995006300020003x

Marquardt, D. W. (1963). An Algorithm for Least-Squares Estimation of Nonlinear Parameters. Journal of the Society for Industrial and Applied Mathematics, 11(2), 431–441. https://doi.org/10.1137/0111030

Martins, J. R. R. A., Sturdza, P., & Alonso, J. J. (2003). The complex-step derivative approximation. ACM Transactions on Mathematical Software, 29(3), 245–262. https://doi.org/10.1145/838250.838251

Mualem, Y. (1976). A new model for predicting the hydraulic conductivity of unsaturated porous media. Water Resources Research, 12(3), 513–522. https://doi.org/10.1029/WR012i003p00513

Pachepsky, Y. A., Guber, A. K., Yakirevich, A. M., McKee, L., Cady, R. E., & Nicholson, T. J. (2014). Scaling and Pedotransfer in Numerical Simulations of Flow and Transport in Soils. Vadose Zone Journal, 13(12), vzj2014.02.0020. https://doi.org/10.2136/vzj2014.02.0020

Peters, A. (2013). Simple consistent models for water retention and hydraulic conductivity in the complete moisture range. Water Resources Research, 49(10), 6765–6780. https://doi.org/10.1002/wrcr.20548

Peters, A., & Durner, W. (2015). SHYPFIT 2.0 User's Manual. Research Report. Institut Für Ökologie, Technische Universität Berlin, Germany. http://www.soil.tu-bs.de/mitarbeiter/durner/software/shypfit2.0/User%20Manual_SHYPFIT%202.0.pdf

Pinheiro, E. A. R., & de Jong van Lier, Q. (2021). Propagation of uncertainty of soil hydraulic parameterization in the prediction of water balance components: A stochastic analysis in kaolinitic clay soils. Geoderma, 388, 114910. https://doi.org/10.1016/j.geoderma.2020.114910

Press, W. H. (Org.). (2007). Numerical recipes: The art of scientific computing (3rd ed). Cambridge University Press.

Squire, W., & Trapp, G. (1998). Using Complex Variables to Estimate Derivatives of Real Functions. SIAM Review, 40(1), 110–112. https://doi.org/10.1137/S003614459631241X

Van Genuchten, M. Th. (1980). A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils. Soil Science Society of America Journal, 44(5), 892–892. https://doi.org/10.2136/sssaj1980.03615995004400050002x

Van Genuchten, M. Th., Leij, F. J., & Yates, S. R. (1991). The RETC Code for Quantifying the Hydraulic Functions of Unsaturated Soils. United States Environmental Reseach Laboratory, December, 93–93. https://doi.org/10.1002/9781118616871

Van Genuchten, M. Th., & Nielsen, D. R. (1985). On describing and predicting the hydraulic properties of unsaturated soils. Annales Geophysicae, 3(5), 615–628.

Wen, X.-H., & Gómez-Hernández, J. J. (1996). Upscaling hydraulic conductivities in heterogeneous media: An overview. Journal of Hydrology, 183(1–2), ix–xxxii. https://doi.org/10.1016/S0022-1694(96)80030-8

Wesseling, J., Kroes, J., Campos Oliveira, T., & Damiano, F. (2020a). The impact of sensitivity and uncertainty of soil physical parameters on the terms of the water balance: Some case studies with default R packages. Part I: Theory, methods and case descriptions. Computers and Electronics in Agriculture, 170, 105054. https://doi.org/10.1016/j.compag.2019.105054

Wesseling, J., Kroes, J., Campos Oliveira, T., & Damiano, F. (2020b). The impact of sensitivity and uncertainty of soil physical parameters on the terms of the water balance: Some case studies with default R packages. Part II: Results and discussion. Computers and Electronics in Agriculture, 170, 105072. https://doi.org/10.1016/j.compag.2019.105072

4. CONCLUDING REMARKS

In this thesis, three main chapters were presented with subjects dealing with the improvement of the prediction and understanding of the soil water retention function (RF) and the hydraulic conductivity function (CF). The RF and CF are keystones to soil hydraulic behavior which can be evaluated by hydrological models. Numerical hydrological models based on the Richards equation are highly robust nowadays, but their output is as reliable as the parameterization of the hydraulic properties used as model input.

Chapter I extends the use of the RF and CF to the drier range where the common models cannot predict retention and conductivity reliably. In Chapter II, a modification of the assumptions for the simplified evaporation method used to parameterize the RF and the CF is proposed. This modification increases the accuracy of the predicted retention and conductivity data, using this widely used laboratory technique. In Chapter III, a parameter transformation to improve the estimation of RF parameter uncertainties resulting from the fitting procedure is proposed. The transformation makes the parameter probability distribution resemble more to a normal distribution, providing a more accurate value for the standard deviation of the parameters and facilitating the application of stochastic techniques.

Overall, the author is confident that the presented innovations may contribute to an improved determination of soil hydraulic properties, increasing accuracy and decreasing uncertainty in parameters and in simulation results obtained by hydrological modeling.