Key Points:
- A pedotransfer function is established to relate van Genuchten-Mualem (VGM) to Brunswick (BW) soil hydraulic property model parameters
- The BW model overcomes the structural deficiencies in VGM and leads to considerably better descriptions of retention and conductivity data
- With the new pedotransfer function, soil properties and information in soil maps can be used to predicted BW model parameters

Correspondence to:
T. K. D. Weber,
tobias.weber@uni-hohenheim.de

Pedotransfer Function for the Brunswick Soil Hydraulic Property Model and Comparison to the van Genuchten-Mualem Model

Tobias K. D. Weber[1] (ID), **Michael Finkel[2,3]** (ID), **Maria da Conceição Gonçalves[3]** (ID), **Harry Vereecken[4]** (ID), **and Efstathios Diamantopoulos[5]** (ID)

[1]Institute of Soil Science and Land Evaluation, University of Hohenheim, Stuttgart, Germany, [2]Center for Applied Geoscience, University of Tübingen, Tübingen, [3]SAFSV-Laboratório de Solos, Instituto Nacional de Investigação Agrária e Veterinária, Oeiras, Portugal, [4]Agrosphere Institute (IBG-3), Forschungszentrum Jülich, Jülich, Germany, [5]Department of Plant and Environmental Science, Section of Environmental Chemistry and Physics, University of Copenhagen, Copenhagen, Denmark

Abstract Modeling soil hydraulic properties requires an effective representation of capillary and noncapillary storage and conductivity. This is made possible by using physically comprehensive yet flexible soil hydraulic property models. Such a model (Brunswick [BW] model) was introduced by Weber et al. (2019, https://doi.org/10.1029/2018WR024584), and it overcomes some core deficiencies present in the widely used van Genuchten-Mualem (VGM) model. We first compared the performance of the BW model to that of the VGM model in its ability to describe water retention and hydraulic conductivity data on a set of measurements of 402 soil samples with textures covering the entire range of classes. Second, we developed a simple transfer function to predict BW parameters based on VGM parameters. Combined with our new function, any existing pedotransfer function for the prediction of the VGM parameters can be extended to predict BW model parameters. Based on information criteria, the smaller variance of the residuals, and a 40% reduction in mean absolute error in the hydraulic conductivity over all samples, the BW model clearly outperforms VGM. This is possible as the BW model explicitly accounts for hydraulic properties of dry soils. With the new pedotransfer function developed in this study, better descriptions of water retention and hydraulic conductivities are possible. We are convinced that this will strengthen the utility of the new model and enable improved field-scale simulations, climate change impact assessments on water, energy and nutrient fluxes, as well as crop productivity in agroecosystems by soil-crop and land-surface modeling. The models and the pedotransfer function are included in an R package spsh (https://cran.r-project.org/package=spsh).

Plain Language Summary Soil hydraulic property models are mathematical functions, which describe the relationship between the soil water pressure head and the state of soil water saturation, on the one hand, and the soil water pressure head and the unsaturated soil hydraulic conductivity, on the other. These types of mathematical functions are flexible by adjustable parameters. With one set of model equations, the hydraulic properties of soils which may have very different properties due to their vast natural variability can be described. The models treated in this work are (i) the van Genuchten-Mualem model, a model with well-known problems, but still frequently applied, and (ii) a relatively new physical comprehensive model, named the Brunswick model. First of all, in a data-based comparison of model performance, we demonstrate that the Brunswick model has systematic advantages. Second, knowledge about these above-mentioned parameters can be determined through other mathematical functions, so-called hydro-pedotransfer functions, which empirically relate these parameters to observed soil properties. The information about these soil properties can be measured in the laboratory and is also recorded in soil maps. We created a new pedotransfer function to facilitate the prediction of model parameters for the new Brunswick model.

1. Introduction

The accurate representation of unsaturated water fluxes in soils is important for reliable descriptions of water and energy fluxes (van Genuchten & Pachepsky, 2011) at the soil-atmosphere boundary (Vereecken

et al., 2019), for the prediction of plant growth and yields (Angulo et al., 2014; Gayler et al., 2009; Hoffmann et al., 2016; Lawless et al., 2008), including irrigation optimization (Elmaloglou et al., 2010), and for the quantification of the environmental fate of agrochemicals (Diamantopoulos et al., 2017; Vereecken et al., 2016). For this, the Richards equation, which is now known as the Richardson equation (Raats & Knight, 2018), is considered as de facto standard (Diamantopoulos & Durner, 2012). As necessary input, it requires functions to describe the soil hydraulic properties (SHPs), of which the most widely used is arguably the van Genuchten-Mualem (VGM; van Genuchten, 1980) model. However, a number of well-known short-comings are summarized by Weber et al. (2019), motivating the authors to present a model framework. In the Weber et al. (2019) model, termed Brunswick (BW) model, the soil water retention curve (WRC) and the hydraulic conductivity curve (HCC) are ramified into a capillary and noncapillary pore space. This is achieved by introducing only one additional model parameter. The noncapillary part accounts for some underrepresented effects such as corner and film flow (Diamantopoulos & Durner, 2013; Tuller & Or, 2001). With the BW model, a water content of 0 at oven dryness is ensured. Further, the inclusion of the noncapillary part enables a linear reduction of the water content with increasing pF, as has been often observed in experimental data (Schelle et al., 2013). Moreover, the often observed change in slope in the mid-pF range in the HCC (Diamantopoulos & Durner, 2015) can also be modeled with great accuracy. Finally, it should be noted that the BW model can be used with any capillary saturation function, which can be integrated in pF space (Streck & Weber, 2020). We define pF by $pF = \log_{10} |h|$, that is, as the common logarithm of the absolute value of pressure head h (cm).

There are three principle methods of gaining knowledge about the SHP: first, by obtaining knowledge of the WRC and HCC from laboratory experiments (van Genuchten et al., 1997) and subsequent parameter estimation; second, through inverse modeling of time series of the relevant state variables (Romano et al., 2002); and third, by pedotransfer functions (PTFs) (van Looy et al., 2017). In general, PTFs predict any difficult to be measured soil property from easily measurable soil properties. When considering soil hydrology, we should, in fact, explicitly refer to hydro-pedotransfer functions (hyPTFs), since other types of PTF exist, too (Goncalves et al., 2001; Martin et al., 2009; Selim, 1998). hyPTFs relate soil properties or soil texture class (and, if available, bulk density, soil organic carbon or matter content, cation exchange capacity, etc.) to SHP and SHP model parameters (e.g., Meskini-Vishkaee et al., 2014; Schaap & Leij, 2000; Tóth et al., 2015). While the first two methods may present more accurate knowledge, they are typically labor intensive, as well as being location and scale dependent. An advantage of hyPTF is that for field-scale application soil texture information may be derived directly from soil maps (Hengl et al., 2017; Tóth et al., 2017). Laboratory measurements can be effective in estimating large-scale effective SHP (Hopmans et al., 2002), but, in spite of the large number of data sets of laboratory measured SHPs included in the construction of hyPTF, concerns linked to scale and effectiveness may certainly prevail (Philip, 1980; Vogel, 2019). We can, nevertheless, expect that if the uncertainties during PTF development are contained (input uncertainty) and reported (output uncertainty), the application of hyPTF to the field scale can capture some problems of scale if adequate approaches (e.g., stochastic modeling; Beyer et al., 2009; Famiglietti & Wood, 1994) are used to represent the soil spatial variability (Hopmans et al., 2002).

The abundant literature on hyPTF contains a multitude of functional relationships between soil texture (properties) and SHP or SHP model parameters (Vereecken et al., 2010). Due to their widespread use in earth system models such as Hydrus 1D (Šimůnek et al., 2016), land atmosphere models (Vereecken et al., 2019), or soil-vegetation-atmosphere models such as the Daisy model (Abrahamsen & Hansen, 2000) and the agroecosystem multimodel library Expert-N (Priesack, 2006), there has been a strong emphasis on VGM model parameters. Some selected studies exist for other models, too (van Looy et al., 2017). Again, it has been demonstrated numerous times that VGM and VGM-type parameterizations of the SHP suffer from well-known problems (Peters, 2013; Weber et al., 2019). However, no readily available hyPTF predicting parameters of more physically comprehensive models such as th*e* BW model exist. In part, this might be due to the fact that the necessary large data sets with measurements of the WRC and HCC are typically not available to the wider scientific community, in spite of the *proviso* of replicable science.

Nonetheless, we are determined to make use of the fact that many numerical codes with implemented and tested hyPTF for the VGM model already exist. We can exploit this fact with the following approach: based on a data set of WRC and HCC data, compiled from literature sources, we established a constitutive
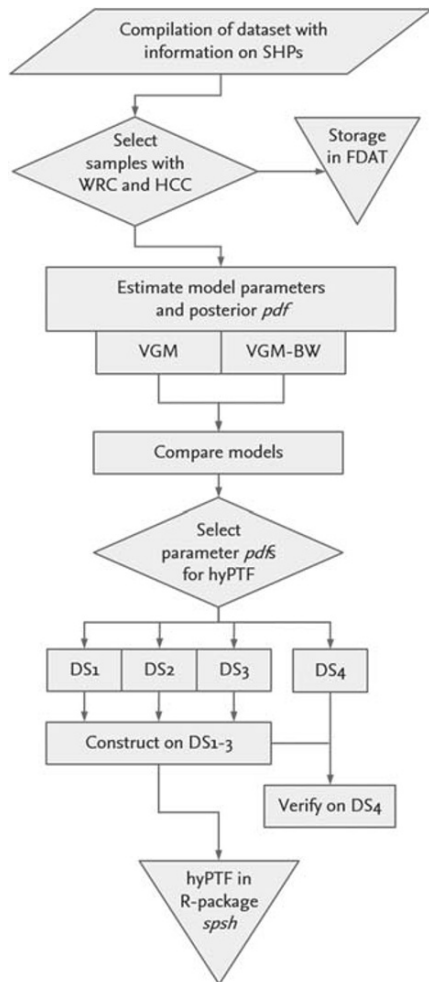
Figure 1. Flowchart of the research approach. The data sets (DS) 1–4 are described in the text.

relationship between the model parameters of the VGM model and the BW model in the VGM variant (BW-VGM). With this strategy, the direct use of existing codes is possible, while enabling a more physically comprehensive representation of the SHP. This leads us to the first hypothesis in which we state that there is a quantifiable relationship between the VGM parameters and the BW-VGM model parameters. Second, we hypothesize that the BW-VGM model, also in a data-based evaluation, is better than the VGM model. Therefore, our aims were to (i) compare the model performance based on a large data set, (ii) establish a parameter equivalence (Assouline & Or, 2013) between the VGM and the BW-VGM model parameters, and (iii) increase the utility of existing hyPTF by extending them to the BW-VGM.

2. Material and Methods

To achieve these aims, the following research approach was adopted, and a flowchart of the research approach is given in Figure 1.

1. Measured WRC and HCC from the literature were compiled in a data set, publicly available in the context of this work at the research data portal of the University of Tübingen, FDAT (https://fdat.escience.uni-tuebingen.de/portal/).
2. Knowledge about the VGM and VGM-BW model parameters and their uncertainties was inferred by inverse modeling assuming identically and independently distributed residuals, the statistical properties of which were approximated by two separate Gaussian distributions: one for the WRC and one for the HCC.
3. The variance of the residual is unknown *a priori*. This is a direct consequence of the lack in information provided in the measurement error of the soil hydrological experiments and the unknown nature of the model error. Therefore, the variances of the combined error (residual, i.e., model + measured error) of the HCC and WRC were estimated as extra nuisance parameters. The estimate of the variance of the error can then be directly used to compare the models with each other.
4. The estimation was carried out in a Bayesian framework so that the uncertainty in the resulting model parameters is proportional to the variance of the likelihoods. The posteriors of the model parameters were sampled using a Markov Chain-Monte Carlo (MCMC) algorithm.
5. The functional relationship between corresponding VGM and BW-VGM model parameters was established for all parameters but one using weighted linear regression with individual weights in both the dependent and the independent variables. These weights were determined directly as the variance of the respective sample of the posterior.

2.1. SHP Models

Here, the two SHP models for the pressure head dependent WRC $\theta(h)$ and HCC $K(h)$ are described. Some of the parameters exist in both models; we have chosen to use the traditional symbols but later in the study differentiate them by adding appropriate suffixes. Details of the robust inverse modeling scheme are given, and summary information describing the compiled data set is presented.

2.1.1. The VGM Model
Soil Water Retention Function

The soil water retention θ ($L^3 L^{-3}$) is a function of pressure head h (L) and given by the van Genuchten equation

$$\theta(h) = \theta_r + (\theta_s - \theta_r)\Gamma(h), \tag{1}$$

where θ_r ($L^3 L^{-3}$) and θ_s ($L^3 L^{-3}$) are the residual and the saturated water contents, respectively, and the effective saturation function $\Gamma(h)$ is given as (van Genuchten, 1980)

$$\Gamma(h) = [1 + (\alpha|h|)^n]^{-m}, \tag{2}$$

where α (L^{-1}), n ($-$), and m ($-$) are shape parameters, with the constraint of $m = 1 - 1/n$.

Hydraulic conductivity curve

The closed-form expression for the HCC $K(h)$ ($L T^{-1}$) is given by

$$K(h) = K_s\Gamma(h)^\tau \left(1 - \left(1 - \Gamma^{(1/m)}\right)^m\right)^2, \tag{3}$$

where K_s ($L T^{-1}$) is the saturated hydraulic conductivity and τ ($-$) is a shape parameter. This leads to six adjustable model parameters which we will denote here as $\mathbf{x_{VGM}} = \{\theta_r, \theta_s, \alpha, n, \tau, K_s\}$.

2.1.2. The BW Model in the VGM Variant

In comparison with the VGM model, the BW framework model conceptually adopts the method of partitioning the SHP, introduced by Peters (2013), into a capillary and a noncapillary part (Weber et al., 2019). The obtained model for the WRC is continuously differentiable in h. In contrast to most other retention functions, a water content of zero at very high-pressure heads $|h|$ is ensured. Finally, the desirable feature of a near linear decrease in water content on a log-linear scale at high-pressure heads $|h|$ is obtained. The capillary and noncapillary conductivities are computed from the respective saturation functions, thereby accounting for more flow phenomena in the drier part of the HCC. All of this is achieved by one additional model parameter. Here, the BW model framework is briefly explained. Full details can be found in Weber et al. (2019).

Soil water retention function

Again, the WRC is a function of h but given as a weighted sum of the capillary $S_c(h)$, and noncapillary $S_{nc}(h)$ saturation function, respectively,

$$\theta(h) = \theta_{cs}S_c(h) + \theta_{ncs}S_{nc}(h), \tag{4}$$

where θ_{cs} ($L^3 L^{-3}$) and θ_{ncs} ($L^3 L^{-3}$) are, respectively, the saturated water content of the capillary and the noncapillary part. In equivalence to Equation 1, the saturated water content, θ_s ($L^3 L^{-3}$) of the VGM model corresponds to the sum of θ_{cs} and θ_{ncs}.

The effective saturation function $S_c(h)$ in Equation 4 is given by rescaling $\Gamma(h)$ (Equation 2) as (Iden & Durner, 2014)

$$S_c(h) = \frac{\Gamma(h) - \Gamma_0(h_0)}{1 - \Gamma_0(h_0)}, \tag{5}$$

where the pressure head value at which oven dryness is attained ($\theta = 0$) is set to $h_0 = -10^{6.8}$ cm, following the suggestion by Schneider and Goss (2012) and in agreement with literature values (Weber et al., 2017b). Other authors suggest to use $h_0 = -10^{6.9}$ cm as anchoring point (Groenevelt & Grant, 2004). Soil water retention values in the very dry end may be determined experimentally with the dew-point method (Gee et al., 1992).

The effective saturation function of the noncapillary part $S_{nc}(h)$ is expressed directly as

$$S_{nc}^*(h) = \log_{10}(e) \int_h^{-10^\epsilon} \frac{S_c(h') - 1}{h'} dh', \tag{6}$$

where h' denotes the dummy variable of integration and the upper boundary of the integral -10^ϵ is a pressure head value very close to 0. From rescaling the Equation 6, an expression for $S_{nc}(h)$ is given as

$$S_{nc}(h) = 1 - \frac{S_{nc}^*(h)}{S_{nc}^*(h_0)}, \tag{7}$$

which ensures that $S_{nc}(h)$ scales between 0 and 1 when h varies between h_0 and ϵ cm. The expression for

$\Gamma(h)$ can be replaced by any given capillary saturation function, such as those proposed by van Genuchten (van Genuchten, 1980), Brooks-Corey (Brooks & Corey, 1964), Kosugi (Kosugi, 1996), or Fredlund-Xing (Fredlund & Xing, 1994), demonstrating the flexibility of this approach for modeling the noncapillary retention function. For this study, we set $\epsilon = -3$. The integral in Equation 6 was solved numerically in h space as implemented in the R package spsh (Weber, Diamantopoulos, et al., 2020). The solution was derived by numerically integrating using 302 support points on the pressure head interval $[10^{-3}; 10^{6.8}]$ and by adding measured pF values during parameter estimation for each sample individually. An analytical solution derived by Streck and Weber (2020) was derived after completion of this study and used as a benchmark for numerical solution. The results showed that no relevant differences exist between the two.

Hydraulic conductivity curve

Similar to the WRC, the liquid HCC is given as a sum of its capillary $K_{\rm c}(h)$ (L T^{-1}) and noncapillary $K_{\rm nc}(h)$ (L T^{-1}) conductivities. Under consideration of the isothermal vapor conductivity $K_{\rm ivc}(h)$ (L T^{-1}), the HCC $K(h)$ is given similarly to Peters (2013) by

$$K(h) = K_{\rm c}(h) + K_{\rm nc}(h) + K_{\rm ivc}(h), \tag{8}$$

where the HCC for the capillary part $K_{\rm c}(S_{\rm c}(h))$ is given by (Peters, 2014)

$$K_{\rm c}(S_{\rm c}) = K_{\rm sc} S_{\rm c}^{\tau} \left[1 - \left(\frac{1 - \Gamma^{\frac{1}{m}}}{1 - \Gamma_0^{\frac{1}{m}}} \right) \right]^2, \tag{9}$$

where the functional argument h has been dropped from the notation and $K_{\rm sc}$ is the saturated capillary conductivity (L T^{-1}).

The HCC for the noncapillary part $K_{\rm nc}(S_{\rm nc}(h))$ is given by Weber et al. (2019) as

$$K_{\rm nc}(S_{\rm nc}) = K_{\rm snc} \left(\frac{h_0}{h_{\rm r}} \right)^{-a(1 - S_{\rm nc})}, \tag{10}$$

where $h_{\rm r} = 1$ cm ensuring matching dimensions, a ($-$) is a parameter, which governs the slope of the HCC in the part of the function, where noncapillary flow dominates and is set to 1.5 which is in agreement with direct measurements (Tokunaga, 2009) and as used in the literature (Weber et al., 2017a). The tortuosity model in $K_{\rm ivc}$ we used was the Millington and Quirk model (Millington & Quirk, 1961). We give $K_{\rm ivc}$ here for completeness but refer to Saito et al. (2006) or Peters (2013) for a full description.

This leads to seven adjustable model parameters denoted by $\boldsymbol{x_{\rm BW-VGM}} = \{\theta_{\rm snc}, \theta_{\rm sc}, \alpha, n, \tau, K_{\rm cs}, K_{\rm ncs}\}$. The shapes of the WRC and HCC of both models are shown in Figure 2, along with the specific soil water capacity function.

2.2. Data Set

A data set of 1,729 samples with WRC and HCC data was comprised from the following data sets: The Portuguese and German data sets from Weynants et al. (2013) (apart from the the BGR data, which we did not use as it was not freely available), the data in the supporting information of Rudiyanto et al. (2015), the UNSODA database from Nemes et al. (2001), and the "Vereecken database" from Vereecken et al. (1989) (Vereecken et al., 2017). The textural data are summarized in Figure 3 showing a coverage not untypical for studies on PTFs and the common underrepresentation of clay soils. The data were compiled and can be accessed publicly (Weber, Vereecken, et al., 2020).

Figure 4 summarizes the data set by reporting the fraction of samples for which measurements above and below certain pressure heads were reported. For example, 90% of the samples contained WRC data with $pF \leq 0.8$ and 50% with HCC data (blue circles). In general, the wet end has more samples with measured WRC than HCC. In the dry end, this is reversed at $pF \geq 4.0$. This is related to the adopted measurement techniques.
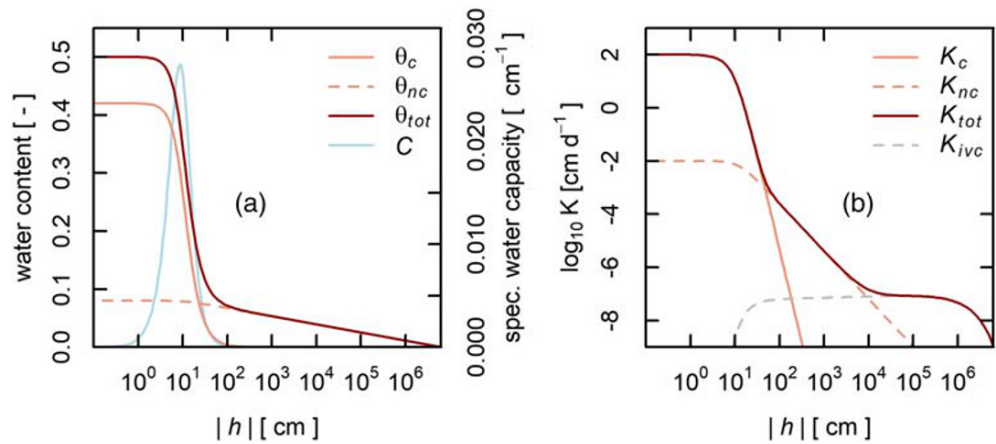
Figure 2. The SHP models (a) WRC and (b) HCC with the BW-VGM parameters $\theta_{snc} = 0.08$ (−), $\theta_{sc} = 0.42$ (−), $\alpha = 0.05$ cm^{-1}, $n = 1.6$ (−), $K_{sc} = 100$ cm d^{-1}, $K_{snc} = 10^{-4}$ cm d^{-1}, and $\tau = 0.5$ (−). Explanation of the parameters is given in the text. The continuous salmon-colored lines are identical to the VGM model, but the retention curve runs between 0 and θ_{sc} (i.e., $\theta_s - \theta_r$) and not θ_s and θ_r.

2.3. Parameter Estimation

In the following, two different approaches for the inference about the parameter are described. First, the method to estimate the SHP model parameters from Bayesian inference is described, in which we employ a converged MCMC sampler, to sample from the posterior distribution of the model parameters. Later, the functional relationships relating the parameters $x_{\mathbf{VGM}}$ to selected parameters in $x_{\mathbf{BW-VGM}}$ are described. Some parameters were transformed for the entire analyses as follows: $\log_{10}(n - 1)$, $\log_{10}(\alpha)$, $\log_{10}(K_s)$, $\log_{10}(K_{sc})$, and $\log_{10}(K_{snc})$. This means the prior and parameter bounds were specified accordingly.

2.3.1. Soil Hydraulic Properties

According to the continuous case of Bayes' theorem, the posterior probability density function (pdf) of the model parameters is (Box & Tiao, 1992)
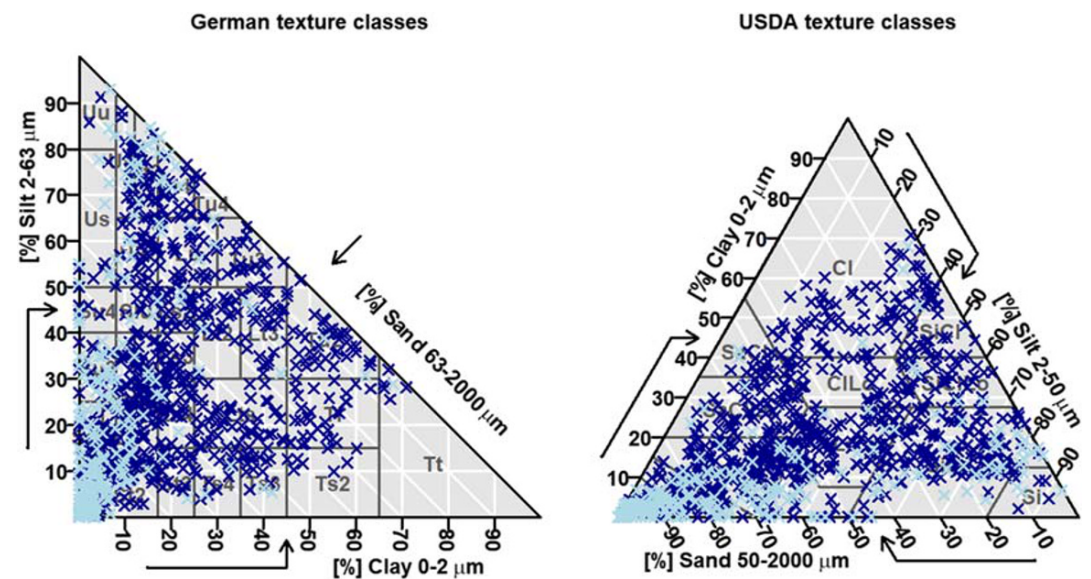
Figure 3. Texture information samples in the database (left) German soil classification system (right) USDA classification. The limits of the texture ranges can be found in the respective subfigures; in dark blue, the entire data set is shown, and in light blue, the reduced data set, DS1 (explanation in section 3.1).
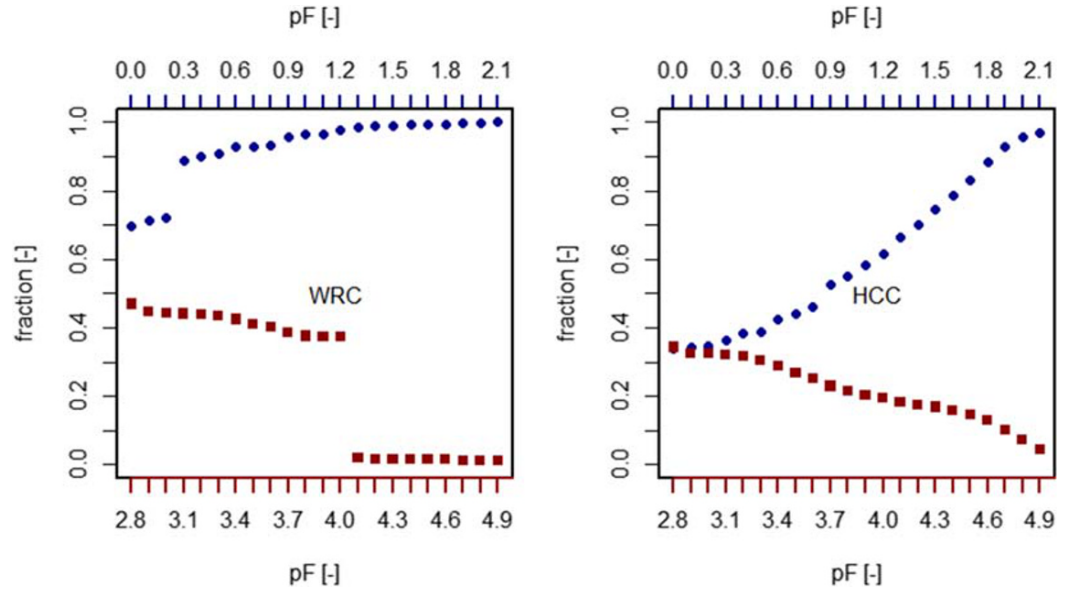
Figure 4. Cumulative frequencies describing the percentage of samples with measurements above (blue circles; top axes) and below (red squares; bottom axes) a certain pressure head ($pF = \log_{10} | h|$), (left) $\theta(h)$ and (right) $\log_{10}K(h)$.

$$f_{\text{post}}(\boldsymbol{x_k}|\boldsymbol{y}) = \frac{L(\boldsymbol{y}|\boldsymbol{x_k})f_{\text{prior}}(\boldsymbol{x_k})}{f(\boldsymbol{y})},$$ (11)

where $f_{\text{post}}(\boldsymbol{x_k}|\boldsymbol{y})$ is the posterior distribution of the model parameter set x_k (i.e., $\boldsymbol{x_{\text{VGM}}}$ or $\boldsymbol{x_{\text{BW-VGM}}}$) conditional on the measured WRC and HCC data set, denoted by $\boldsymbol{y}$. The term in the denominator of Equation 11, $f(\boldsymbol{y})$, is independent of the model parameters and serves as a normalizing constant so that $f_{\text{post}}(\boldsymbol{x_k}|\boldsymbol{y})$ integrates to unity. We explicitly note that, in formulating Equation 11, we do not explicitly formulate prior model weights (Schöniger et al., 2014). However, we do note that through the different number of parameters, different $f_{\text{prior}}(\boldsymbol{x_{\text{VGM}}})$ values are calculated, which are not equal for VGM and BW-VGM. This is an indirect result of the latter model carrying an additional parameter (K_{snc}), while the parameter bounds which correspond in both models were not altered. Scharnagl et al. (2011) demonstrated how the choice of prior may considerably influence the shape and the mode of $f_{\text{post}}(\boldsymbol{x_k}|\boldsymbol{y})$ for SHP parameters, but in the interest of a fair comparison, the prior was bounded, but no further shape assumptions beyond bounded uniform were included. Therefore, we are assuming here that $f_{\text{prior}}(\boldsymbol{x_k})$ is best described by a multivariate distribution with uniform distributed bounded marginals; the bounds are presented in Table 1. From this, we can compute $\ln f_{\text{prior}}(\boldsymbol{x_{\text{VGM}}}) = -3.51$ and $\ln f_{\text{prior}}(\boldsymbol{x_{\text{BW-VGM}}}) = -5.3$, implying prior weights of 6:1 in favor of the VGM model. The assumptions on the prior, here, render the posterior proportional to the likelihood function. The posterior sampling of $f_{\text{post}}(\boldsymbol{x_k}|\boldsymbol{y})$ by a MCMC is only known to a constant of proportionality (Gelman et al., 2014) so that MCMC sampling is possible without specifying $f(\boldsymbol{y})$. The joint likelihood to observe the data is then given by

$$L(\mathbf{y}|\mathbf{x_k}) = \prod_{i=1}^{N_\theta} L_{i,\,\theta}(\mathbf{y_{i,\theta}}|\mathbf{x_k}) \times \prod_{i=N_\theta+1}^{N_\theta+N_{\log_{10}K}} L_{i,\,\log_{10}K}(\mathbf{y_{i,\log_{10}K}}|\mathbf{x_k}),$$ (12)

where N_θ is the number of data points in the observed WRC, $N_{\log_{10}K}$ is the number of the observed HCC, and $y_{i,\theta}$ and $y_{i,\,\log_{10}K}$ are the ith measurements in the WRC and HCC data set, that is, the different data groups.

We point out that we have simplified the notation and subsequently dropped an explicit reference to $L_{i,\theta}$ and $L_{i,\,\log_{10}K}$. In the following we assume statistically independent and heteroscedastic errors (homoscedastic within θ and K) which are calculated by

Table 1
Bounds for the Shp Model Parameters of the Prior, and the Two Nuisance Parameters

Parameter	Unit	Lower bound	Upper bound
θ_r , θ_{snc}	(–)	0.001	0.35
θ_s , θ_{sc}	(–)	0.2	0.7
α_{VGM} , α_{BW}	(cm^{-1})	0.001	0.1
n_{VGM} , n_{BW}	(–)	1.1	11
K_S , K_{SC}	(cm d^{-1})	1	1,000
K_{snc}	(cm d^{-1})	1 E-6	1 E-1
τ_{VGM} , τ_{BW}	(–)	−2	10
Statistical model			
σ_θ	(–)	0.001	1
$\sigma_{\log_{10}K}$	(cm d^{-1})	0.001	1

Note. Explanation of the parameters is given in the text. The parameters α, $n - 1$, K, and σ were estimated in the $\log_{10}$-transformed space.

Table 2
Comparison of the Median Mean Error (ME) and Median Mean Absolute Error (MAE) Between the Van Genuchten-mualem (Vgm) and the Brunswick (BW-VGM) Model for the Two Data Groups Water Retention Curve (WRC) and Hydraulic Conductivity Curve on the 402 Samples

	WRC		HCC	
	($cm^3 \, cm^{-3}$)		($\log_{10}$ [cm d^{-1}])	
Model	ME	MAE	ME	MAE
VGM	−1.5 E-8	4.6 E-3	−1.7 E-5	0.17*
BW-VGM	−7.8 E-6	4.6 E-3	2.4 E-5	0.11*

*Denotes significant difference with $p = 1.48$ E-8 between the models ME and MAE for HCC.

$$e_{ik}(h_i, \boldsymbol{x_k}) = y_{ik} - f(h_i, x_k) \qquad k = 1, 2 \quad \text{and} \quad i = 1, ..., N, \qquad (13)$$

where $f(h_i, x_k)$ denotes the model-predicted WRC or HCC value at pressure head h_i and model for k, with $N = N_\theta + N_{\log_{10}K}$. The errors in WRC and HCC are assumed to have a zero mean ($E[e_{ik}(h_i, \boldsymbol{x_k})] = 0$) and are normalized by the respective standard deviations σ_{lk}, where l denotes the data group of WRC and HCC. The standardized residual $\underline{e_{ik}}$ is obtained by

$$\underline{e_{ik}}(h_i, \boldsymbol{x_k}) = \frac{e_{ik}}{\sigma_{lk}} \qquad i = 1, ..., N. \qquad (14)$$

The fact that σ_{lk} can be different for WRC and HCC leads to the choice of estimating an individual σ_{lk}. Thus, we explicitly write σ_θ and $\sigma_{\log_{10}K}$ which were estimated as additional nuisance parameters in the Bayesian inference for each of the two models with the same bounds (Table 2). As illustrated in Weber et al. (2018), the choice of the likelihood model $L(y| \boldsymbol{x_k})$ influences $f_{\text{post}}(\boldsymbol{x_k}| y)$. With reference to the above, this can lead to better approximations of $f_{\text{post}}(\boldsymbol{x_k}| y_i)$ but requires close, that is, supervised examination of the inference result for each individual sample and model. Since this is not within the scope of this study, we assume that $\underline{e_{ik}}$ is approximately normal and the likelihood model for HCC and WRC, under the consideration of Equation 14, is then formulated as

$$L_i(\boldsymbol{y_i}| \boldsymbol{x_k}, \boldsymbol{\sigma_{lk}}) = \frac{1}{\sqrt{2\pi}\sigma_{lk}}\exp\left(-0.5 \, \underline{e_{ik}^2}\right). \qquad (15)$$

For the estimation of the best parameter set, we used a parallelized global optimization algorithm (differential evolution; Storn & Price, 1997) as implemented in the R package DEoptim (Mullen et al., 2011) using a maximum of 500 iterations, default settings, and setting the number of population members as 10 times the number of parameters. The obtained best parameter estimates were used in subsequent analyses. To sample from the posterior, we used the delayed rejection adaptive Metropolis algorithm (Haario et al., 2006) as implemented in the FME package in R (Soetaert & Petzoldt, 2010). The sampling was performed using a single chain with 50,000 iterations. Convergence was attested by ensuring the Gelman and Rubin's convergence diagnostic $\widehat{R}$ (Brooks & Gelman, 1998; Gelman & Rubin, 1992) fulfilled the condition $\widehat{R} < 1.2$, for each element in $\boldsymbol{x}$, calculated based on two additional chains. Of all metrics, the $\widehat{R}$ statistic provides the best guidance on exactly when convergence has been achieved (Vrugt, 2016). The sample was obtained using a maximum number of three tries for the delayed rejection procedure and an adaptation of the covariance matrix of the jump distribution every 100 iterations. The first 90% of the samples were treated as burn-in and discarded, and subsequent analyses were done on the last 10%. As starting point for the chain, we used the best parameter estimate for each given sample and model and added some small amount of noise to avoid the sampler getting stuck at the mode right from the outset.

2.3.2. Construction of the hyPTF
The construction of the hyPTF was based on establishing functional relationships between the estimated parameters (see section 2.3.1) for the two models ($\boldsymbol{x_{\text{VGM}}} \to \boldsymbol{x_{\text{BW-VGM}}}$, except for K_{snc}). We used weighted least squares to account for errors in both the predicted and the explanatory variable (Glaister, 2001) for $\theta_{\text{snc}}(\theta_r)$, $\theta_{\text{sc}}(\theta_s - \theta_{\text{snc}})$, $\alpha_{\text{BW}}(\alpha_{\text{VGM}})$, $n_{\text{BW}}(n_{\text{VGM}})$, $K_{\text{sc}}(K_s)$, and $\tau_{\text{BW}}(\tau_{\text{VGM}})$, while for K_{snc}, we simply used the median of a set of estimated $\widehat{K}_{\text{snc}}$. Next to the adopted weighted linear regression for the remaining parameters, we tested several other regression and machine learning models to predict $\widehat{K}_{\text{snc}}$ based on different parameter combinations. None of the results were considered robust; thus, we suggest using median $\widehat{K}_{\text{snc}}$. The consideration of errors was achieved by using the calculated variance from the posterior's samples $f_{\text{post}}(\boldsymbol{x_k}| \mathbf{y})$ marginal for each parameter. The variance of parameter j in model k for the observed data set d was calculated from the MCMC sample of the posterior and used as a weight $w_{jkd} = [var(\boldsymbol{x_{jkd}})]^{-1/2}$ for the

York regression (York, 1966). We follow this procedure so that those parameters estimated in section 2.3.1 with a large variance, automatically receive a smaller weight in the regression, hereby reducing the biases in the hyPTF. In this way, less well-constrained parameters are assigned lower weights in the regression in contrast to well-constrained parameters.

For estimating unbiased confidence intervals for the regression, we used nonparametric bootstrapping (Davison & Hinkley, 1997) as described in the following: for each relationship between the ith VGM and BW-VGM model parameters, we define a sampled data set $D_i : \left\{ x_j^{\text{VGM}}, x_j^{\text{BW} - \text{VGM}} \right\}$ where i is the index for the parameter (1–6) and j is the index for the $x_{i,j}^{\text{VGM}}$, $x_{i,j}^{\text{BW} - \text{VGM}}$ pairs, determined by MCMC. For each D_i, we repeated the following steps 10,000 times:

1. We generated a bootstrap sample B_i by randomly sampling m points ($m = 1, ..., 10{,}000$) with replacement from D_i with their corresponding variances for both VGM and BW-VGM, as determined by MCMC sampling;
2. we performed a York regression on each B_i and obtained estimated intercepts $\widehat{a}_i$ and slopes $\widehat{b}_i$;
3. then, the set of predictions was calculated $\widehat{x}_n^{\text{BW} - \text{VGM}} : \left\{ \widehat{a}_{ii}^q + \widehat{b}_{ii}^q x_n^{\text{VGM}} \right\}$
4. and calculated the 90% percentile confidence intervals of $\widehat{x}_n^{\text{BW} - \text{VGM}}$, $\widehat{a}_i^q$, and $\widehat{b}_i^q$.

2.4. Model Comparison and Evaluation

For the model comparison, we use several metrics to compare the model performance: (i) the mean error (ME) describing the average bias, (ii) the mean absolute error (MAE) describing the magnitude of the bias, (iii) the root mean squared error (RMSE) describing the spread of the error around the mean, (iv) the standard deviation of the residual for each sample and data group, which can be seen as a contributor to the variance of the model parameters, in the presence of a uniform $f_{\text{prior}}(\boldsymbol{x_k})$, and (v) the Akaike information criterion (AIC; Akaike, 1974) and the Kashyap information criterion (KIC; Kashyap, 1982), metrics rooted in information theory, used as model ranking metrics (Ye et al., 2008). ME and MAE are used as summary statistics over all results.

First, we directly compare σ_θ and $\sigma_{\log_{10} K}$ (as introduced in 2.3.1) as a measure for the scatter of the observed data around the calibrated model. The assumption is that the total variance of the residual $\sigma_{lk, \text{tot}}^2$ in data group (WRC, HCC) and of model k (VGM, BW-VGM) is described by the

$$\sigma_{lk, \text{tot}}^2 = \sigma_{l, \text{OBS}}^2 + \sigma_{lk, \text{MOD}}^2, \tag{16}$$

where $\sigma_{l, \text{OBS}}^2$ is the measurement error and $\sigma_{lk, \text{MOD}}^2$ is the model error, and $\sigma_{lk, \text{tot}}^2$ is the estimated parameter from maximizing Equation 12. In a sample wise comparison of model performance, the σ_{OBS}^2 can be seen as a constant for each individual sample, which is not known. Since it is constant, we can directly compare $\sigma_{l, \text{VGM,tot}}$ with $\sigma_{l, \text{BW,tot}}$ as a qualitative measure for the spread of the residual as a result of the model error of the corresponding model. The smaller value is to be favored.

The RMSE for data group (WRC and HCC) and model k (VGM, BW-VGM) is given by

$$\text{RMSE}_{lk} = \sqrt{\frac{1}{N_l} * \sum_{i}^{N_l} \left[y_{il} - f(h_{il}, x_k) \right]^2}, \tag{17}$$

where N_l is the number of observations in data group l. We acknowledge that there is an equivalence between the comparison of RMSE_{lk} and σ_{lk}.

Lastly, we used two different information criteria to evaluate the model performance by using the concepts behind the AIC and the KIC. According to Höge et al. (2018), AIC is the criterion for identifying the model with the *largest predictive capability*, and KIC is the criteria to identifying which model is *the most likely data generating process*.

The AIC for model k is given by

$$\text{AIC}_k = -2\ln L(\boldsymbol{y}|\widehat{\boldsymbol{x}}_{\boldsymbol{k}}) + 2N_{\widehat{x},k}, \tag{18}$$

where $N_{\widehat{x}}$ is the number of estimated parameters and $L(\boldsymbol{y}|\widehat{\boldsymbol{x}}_{\boldsymbol{k}})$ is the maximum likelihood estimate in Equation 12. The KIC is given by

$$\text{KIC}_k = -2\ln L(\boldsymbol{y}|\boldsymbol{x}_{\boldsymbol{k}}) - 2\ln f_{\text{prior}}(\boldsymbol{x}_{\boldsymbol{k}}) - N_{\widehat{x},k}\ln(2\pi) - \ln\left|\boldsymbol{C}_{\widehat{\boldsymbol{x}\boldsymbol{x}}}\right|, \tag{19}$$

where $\ln f_{\text{prior}}(\boldsymbol{x}_{\boldsymbol{k}})$ is the prior pdf value, and $|\boldsymbol{C}_{\widehat{\boldsymbol{x}\boldsymbol{x}}}|$ is the covariance matrix of the posterior, calculated from the MCMC sample. By comparison, models with lower values are to be favored. Interpreted in model weights, a difference in 10 signifies that the model with the lower value receives 100% of the weight, and the other model may be rejected outright (Burnham & Anderson, 2010).

3. Results and Discussion

3.1. Data Selection

From the parameter estimation of the SHP models, some initial analyses and subsequent sample removal were required, since it appears obvious to construct a hyPTF based only on data which contains sufficient information to constrain model parameters and where the sampler is likely not to have been influenced by the prior bounds. Samples with instances of estimated parameter close to the limits of the already wide prior bounds (Table 1) were removed. This was achieved by removing all samples for which the estimated parameters values did not fulfill the condition $\theta_s < 0.69$, $\tau_{\text{VGM}} < -1.5$, and $\theta_{\text{snc}} < 0.29$. After applying these criteria, the initial data set was reduced to 402 of which 10 samples were chosen at random and set aside for use in the hyPTF model performance (DS4), by way of example. The remaining 392 soil samples compose data set 1 (DS1; Figure 3, light blue crosses). Parameter τ is frequently poorly defined in the VGM model and lacks in physical meaning (Peters et al., 2011), as evident from literature frequently reported as having negative values (e.g., McCarter & Price, 2014; Schelle et al., 2011). In the presence of data from the dryer ranges of the HCC, this parameter can help achieve an improved match of model and data for the frequently observed change in slope in the HCC within the mid-pF range. In other words, in the VGM model, very small τ, that is, unphysical $\tau < 0$, mostly enables a better match of the HCC where noncapillary conductivity becomes visible in the data. Honoring this, the functional parameter relation for τ was based on a reduced data set (DS2), filtering out the samples in DS1 which had $\tau_{\text{VGM}} < 0$. For the hyPTF of K_{snc}, we used DS1 with additional filter criteria discarding $K_{\text{snc}} < 10^{-5}$ cm d^{-1} resulting in DS3 with 359 samples. This precaution was taken to ensure that the MCMC sampler was hardly affected by the lower boundary of the prior. Moreover, the values $<10^{-5}$ showed a large variance in the posterior, indicating that the information in the data was, at best, limited to constrain the parameter.
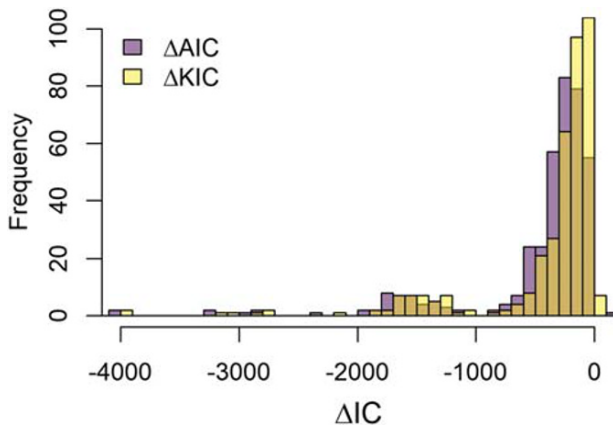
3.2. Model Comparison

In Table 2, the median MEs and MAEs of the models are presented for DS1. The ME in all cases is essentially 0. The only significant difference in model performance could be attested for the MAE in the HCC, where the BW-VGM model was found to be significantly better than the VGM model (Wilcoxon rank sum test, $p = 1.4$ E-8). The ME and MAE of the WRC for both VGM and BW-VGM are very similar. The MAE for the WRC is so small that in standard laboratory settings, these values would not be differentiable. Caution is called for when interpreting this result. As a reminder, the VGM comes with greater flexibility for the WRC, because it is not forced to match a water content of 0 at $h = 10^{6.8}$, considering that most of the samples contain data only up to a pF of about 3 (Figure 4). In other words, particularly for data sets containing observations only from limited pressure head ranges (Scharnagl et al., 2011), the VGM can match the data well. So from a utilitarian point of view, it is treated as an accepted model (Fatichi et al., 2020), proven by the myriad of applications it has found. This is a direct consequence of the VGM being more flexible at the cost of not maintaining physical comprehensiveness and is therefore not as rigid as the BW-VGM, which, in turn, is forced always to meet the physically necessary condition of a water content of zero at a fixed pressure around $pF = 6.8$. The VGM can perform well and might even be superior when matching observed data. However, it does this by releasing the model from physical comprehensiveness, so it cannot claim to describe the properties of a semidry or dry soil, as it is the case for the BW-VGM. This points to the fact that the data are a
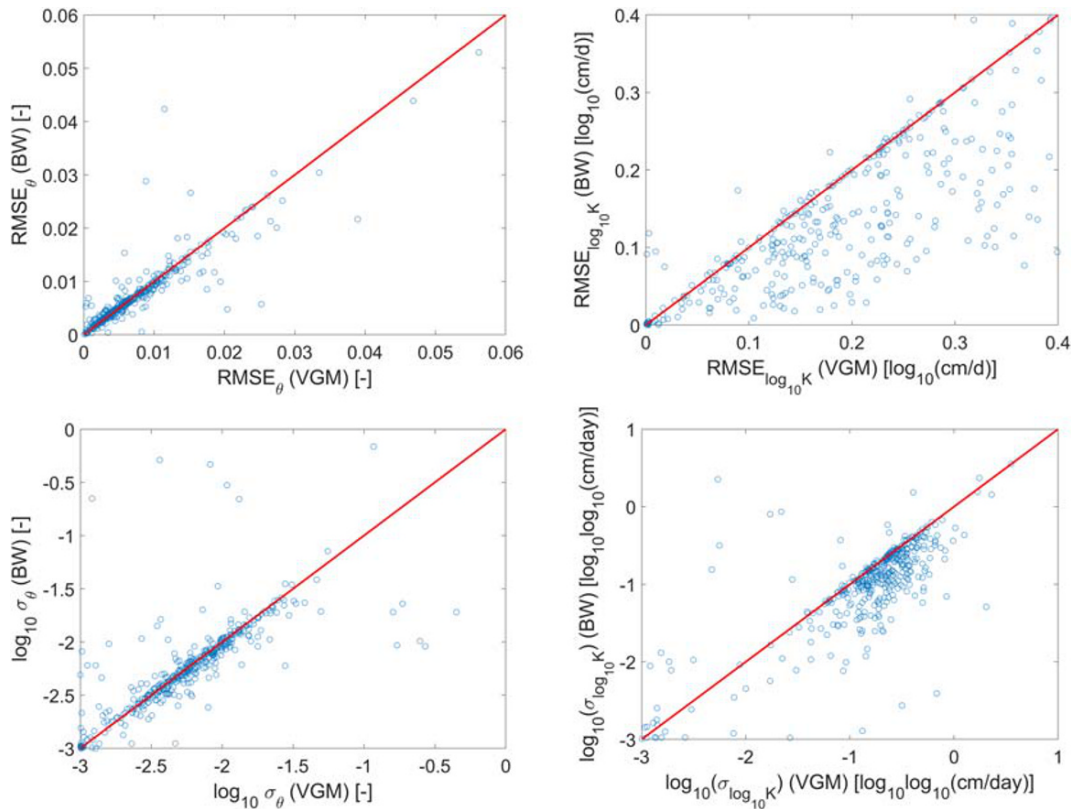
Figure 5. (Top row) Comparison of the estimated combined error variance for $\theta(h)$ (left) and for $\log_{10}K(h)$ (right) between the Brunswick (BW) model and the van Genuchten-Mualem model (VGM). The combined error variance is the lumped model and measurement error variance. (Bottom row) Comparison of the root mean square error (RMSE) for θ (left) and for $\log_{10}K$ (right) between BW and VGM. Red line is the 1:1 line.

limiting factor, since it does not cover the entire pressure head range, highlighting the need for routinely obtaining SHP good data in the dry (and very wet) range of both HCC and WRC. Summarizing our findings, we can state that based on the comparison of the ME and the MAE for the WRC data, the model-data match for BW-VGM is, on average, equal to that of the VGM. On the other hand, for the HCC, the MAE is significantly larger for the VGM, while no significant difference could be distinguished for the ME. This means that on average, both models perform equally well, but the VGM model does not represent hydraulic properties of soils in the dry end of the WRC and HCC. With regard to the overall performance, the magnitude of the bias is significantly reduced by introduction of the physically more comprehensive model, which is mostly an attributed to the better performance in HCC.

This fact is further supported by Figure 5 where the RMSE values and estimated variances σ_θ and $\sigma_{\log_{10}K}$ are shown. In the RMSE of the WRC (Figure 5, top left), a slight scatter around the 1:1 line is observable. This is similar for the estimated variances (Figure 5, bottom left), where, in the majority of cases, the variances are near equal (Figure 5, bottom left). This is different for the HCC (Figure 5, top right panels). Here, almost all estimated RMSE and variances are plotted below the 1:1 line, and the BW-VGM model performs significantly better than VGM (Wilcoxon-Rank-Sum test $p = 1.68$ E-10). This is linked to the VGM model's ability to describe data, which is in a pressure head range attributed to

Figure 6. Difference in the Akaike (AIC) and Kashyap Information Criterion (KIC) between the two models.

the noncapillary part, which can only be achieved by foregoing descriptive capability of data in the capillary part, a direct consequence of the lack in physical comprehensiveness. This is in contrast to the BW-VGM, which can describe this change in slope in the mid-pF range of the HCC (Fig. 2b), expressed in a smaller residual variance.

With regard to the evidence from the Bayesian selection criteria, the above-mentioned better performance is further corroborated by the histograms of the differences between the AIC and KIC (Figure 6), which clearly favor the BW-VGM model. The computed differences, $\Delta KIC = KIC_{\text{BW-VGM}} - KIC_{\text{VGM}}$ and $\Delta AIC = AIC_{\text{BW-VGM}} - AIC_{\text{VGM}}$, are in almost all cases clearly below 10. The ΔAIC values are similar in magnitude to what Weber et al. (2017a) found, in their comparison of performance of the VGM model to an improved SHP model (the PDI model in the van Genuchten variant). The ΔAIC exhibits a slight shift to the left of the ΔKIC indicating that the AIC makes a clearer statement than the KIC. While both criteria are based on the maximum likelihood estimate $-2\ln L(\boldsymbol{y}|\hat{\boldsymbol{x}}_k)$, the penalty terms vary. Both contain a term for the number of parameters, which is multiplied by a factor. This factor is 0.16 larger for the AIC (Equations 18 and 19). Following this, the AIC initially has a larger penalty than the KIC. However, the KIC contains extra penalty for extra parameters when they are not well defined through the prior, in particular if it is noninformative as in our case (Table 2). This difference, expressed in comparative weights, means a 6:1 favoring of the VGM model. This should not be confused with prior model weights; it is the calculated value of the prior at the best parameter estimate. This could have been slightly altered, by reducing the bounds of the θ_{cs} to smaller values, but we chose to use the same bounds as for θ_{s}, as we do not *a priori* know the value of θ_{ncs}. At that point, we had no knowledge of realistic values for θ_{ncs}; it could also be 0, implying the saturated water content could equal the capillary water content $\theta_{\text{s}} \triangleq \theta_{\text{cs}}$. This meant that the bounds were identical, and no further constraint could be legitimized. The introduction of badly constrained parameters is another penalty term, reflected in the term $\ln|C_{\widetilde{xx}}|$. These two factors are larger than the $0.16N_{\widehat{x},k}$ but not large enough to induce a $\Delta KIC > 0$, and a systematic
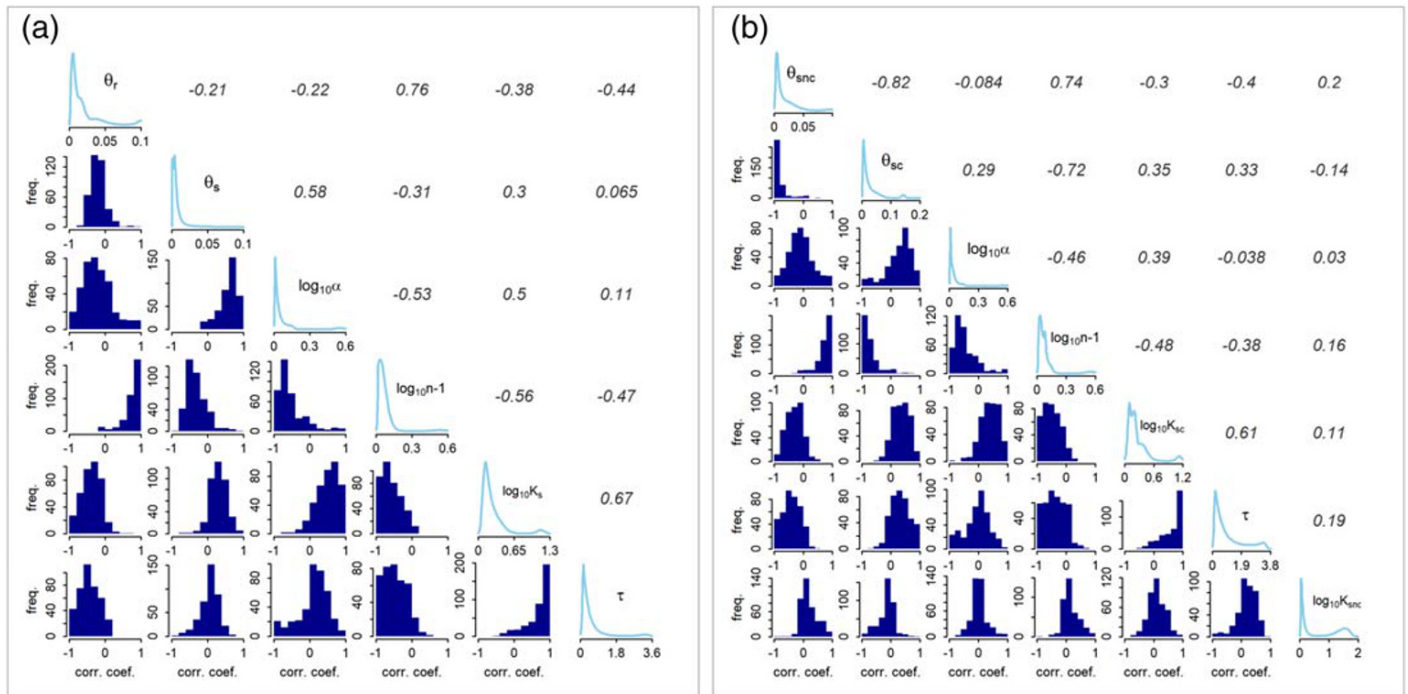
Figure 7. (Lower triangle) Histograms of parameter correlation coefficients between parameters for the (a) VGM model and (b) BW-VGM model, (upper triangle) the corresponding mean correlation coefficients, and (diagonal) densities of the standard deviations of the respective parameter posteriors. Mark, the histograms are unitless, while the units of the densities on the x-axis correspond to those of the respective (transformed) parameters.

favoring of the BW-VGM is the result. This is noteworthy, since the KIC, initially, favors the VGM through the use of the prior, before either the data or parameter dependent terms are added. But from Figure 6, we can read that even if the prior favored the VGM over the BW-VGM 6:1, this is overcome "after the data has been seen by the models" in Equation 12. We note that the AIC does not account for the prior information and thus is also suitable to non-Bayesian methods but can demonstrate the similarity in results, if this study had been conducted using simple weighted least squares.

Information on the robustness of parameter estimates is presented in Figure 7. The lower triangle shows histograms describing the distribution of correlation coefficients between parameters. These were derived by calculating the parameter correlation matrix from the posterior sample of the parameter estimation for each individual soil sample in DS1. Note that for parameter τ, we reduced the data set a little further which does not, however, significantly change the result presented here. Correspondingly, the upper triangle gives the mean correlation coefficient between parameters. The diagonal shows the densities of the standard deviations of all marginals. This was done for both models individually. The density plots indicate small standard deviations and therefore, in almost all cases, clearly identifiable parameters. The mean correlation coefficients are small to moderate. While θ_r and θ_s hardly correlate (−0.21), their counterparts in the BW-VGM correlate more (−0.82). However, the correlations of all other parameters do not appear to be very strong (either close to 1 or close to −1). Overall, no clear picture can be derived as to whether or not correlation increased by adopting the BW-VGM model. This is significant because it means that adopting the physically more comprehensive model together with an additional model parameter does not come at the expense of losing robustness and identifiability.

3.3. hyPTF: Functional Relationship Between Model Parameters
3.3.1. θ_{sc}, θ_{snc}, α, n, K_{sc}, and τ

The predictions of all BW-VGM model parameters, except for K_{snc}, were based on York regression. The corresponding regression parameters (Table 3) can be directly used to derive BW-VGM parameters from VGM parameters. We remark that the intercepts acquire values which are essentially 0. The regression results are depicted in Figure 8, and the corresponding sum of the variance in x and y is given alongside them. First of all, there is a perceivable difference between the weighted line regression and the ordinary least squares (OLS). This can be attributed to the fact that, in York regression, weights are considered leading to robust and unbiased estimates of hyPTF parameters. This is not the case for the saturated water content where OLS and the York estimate are nearly identical as well essentially identical to the 1:1 line. With regard to the physical interpretation of the model parameters, this appears to be consistent, as the modulation in the BW-VGM model affects the dry end and not the total saturated water content of a soil. For the θ_r and θ_{ncs} this is a little different, where for the hyPTF slope for $\theta_{ncs}(\theta_r)$ is 1.28, this means the underestimation of a residual water content, in other words the water content which is not part of the capillary water, is underestimated by about 30% in the VGM model. This may also be an artifact when considering that for many of the samples the data were obtained only from evaporation experiments. In evaporation experiments, the WRC data are usually available until about $pF \sim 3$. This coincides with the pressure head region in which soils over a large range of textures only then show a decrease in the slope of the WRC on the semi-log scale as the soil dries. Exceptions are coarse-grained soils (e.g., sands) with low θ_r, where the capillary water content, in the BW model sense, may have been nearly fully drained by $pF \sim 3$. This means that in nonsandy soils, θ_r often is used to enable a better model-data match, in particular for data in the pressure head range with the steep slope. In other words, θ_r is much more of a fitting parameter than a physically consistent terminology and only delivers limited information about a "residual water content." Moreover, if WRC data do not exist for pressure heads up to $pF > 6$, like in Schelle et al. (2013), then changes in θ_r lead to small changes in the objective function values, meaning that θ_r is less well constrained. Therefore, the estimation of θ_r often leads to larger uncertainties than for the other WRC parameters (e.g., Dettmann & Bechtold, 2016).

In BW-VGM, θ_{snc} is larger (Figure 8) since the matching of the WRC data in the pressure head range of $pF > 3$ is achieved by the near-linear slope of the noncapillary retention curve, which appears only to be achievable when $\theta_{snc} \geq \theta_r$. It is noteworthy that $\log_{10}\alpha$ and $\log_{10}(n − 1)$ are nearly identical. We speculate that this stems from an interplay with τ, K_{sc}, and K_{snc}, and the information content in most observed data sets provides sufficient information content for the parameters of the capillary model to account for the air

Table 3
Regression Coefficients and their Standard Errors (In Brackets) for the Prediction of the BW-VGM Model Parameters from the York Regression

Parameter	Intercept	Slope
θ_{snc} ($\theta_{r,VGM}$)	−1.58 E-03 (7.49 E-4)	1.285 (0.0074)
θ_{tot} ($\theta_{s,VGM}$)	1.89 E-03 (3.00 E-3)	0.993 (0.0072)
$\log_{10}\alpha_{BW}$ ($\log_{10}\alpha_{VGM}$)	−2.06 E-02 (6.30 E-3)	0.986 (0.0036)
$\log_{10}n_{BW} - 1$ ($\log_{10}n_{VGM} - 1$)	6.42 E-02 (2.3 E-3)	0.933 (0.0065)
$\log_{10}K_{s,BW}$ ($\log_{10}K_{s,VGM}$)	1.16 E-01 (1.78 E-2)	1.060 (0.012)
$\tau_{BW}(\tau_{VGM})$ *	2.95 E-02 (1.16 E-1)	1.833 (0.123)
$\log_{10}K_{snc}$**	−1.72**	

*For $\tau_{VGM} \geq 0$, only; for $\tau_{BW}(\tau_{VGM} < 0) = 0$.   **For K_{snc}, the median of the fitted value as indicated is directly used.

entry pressure head and the width in the unimodal pore-size distribution. Generally, the larger τ, the greater the steepness of the capillary bundle model describing the HCC. In Figure 6 (left), it can be seen that the slope of the York fit for τ is so large that hardly any negative $\tau_{BW - VGM}$ remains. This is physically consistent (Peters, 2013) and an indicator that through inclusion of the noncapillary conductivity model part, τ is no longer used to account for HCC in the dry range, which can be achieved
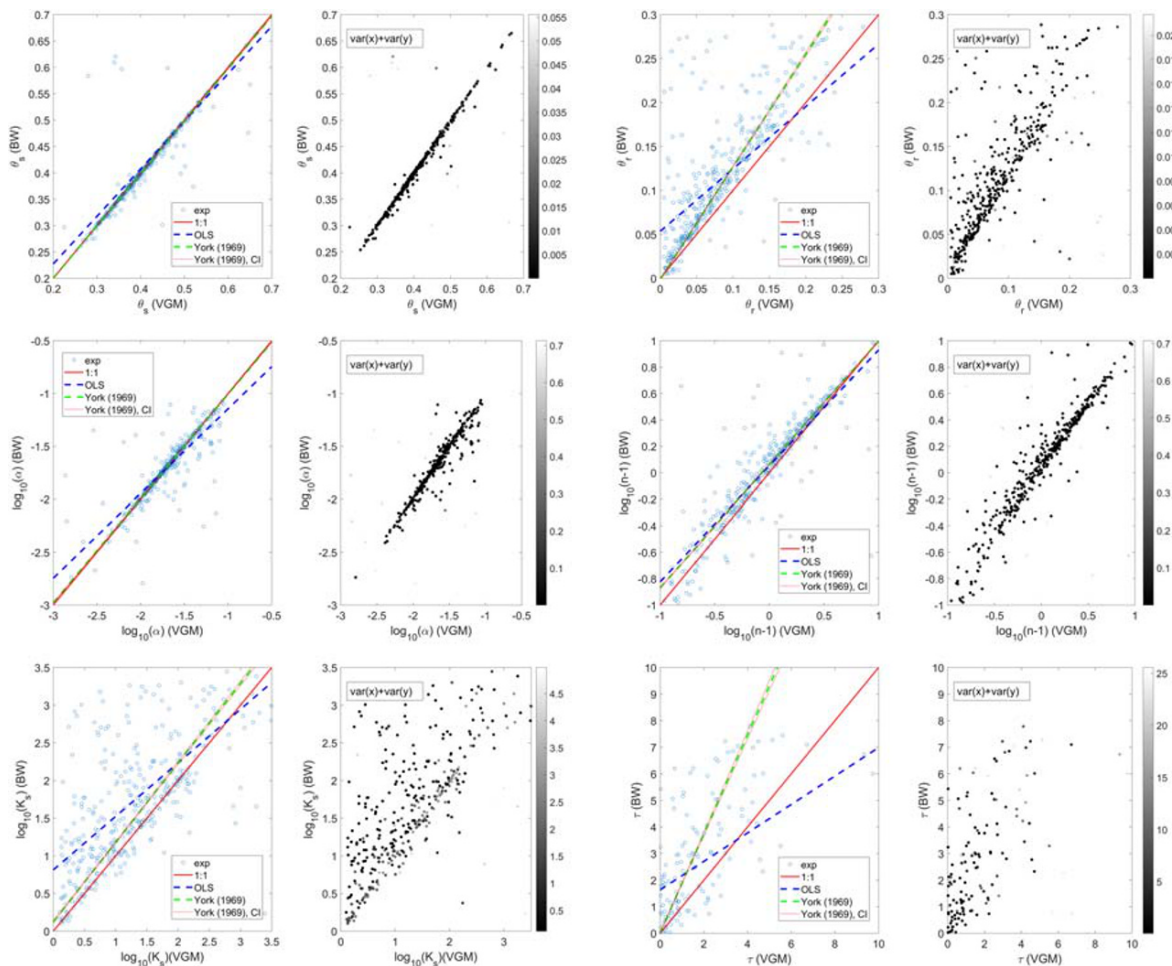
Figure 8. Relationship between the soil hydraulic property model parameters from top to bottom θ_r: θ_{snc} (here denoted as θ_r), θ_s: ($\theta_{sc}+\theta_{snc}$), α_{VGM}: α_{BW-VGM}, n_{VGM}: n_{BW-VGM}, K_s: K_{sc}, τ_{VGM}: τ_{BW-VGM}, τ_{VGM}: τ_{BW-VGM}, constrained (α_i, $n_i - 1$, and K_i are shown as $\log_{10}$-transformed values on both axes), OLS: ordinary least squares, CI: confidence interval.
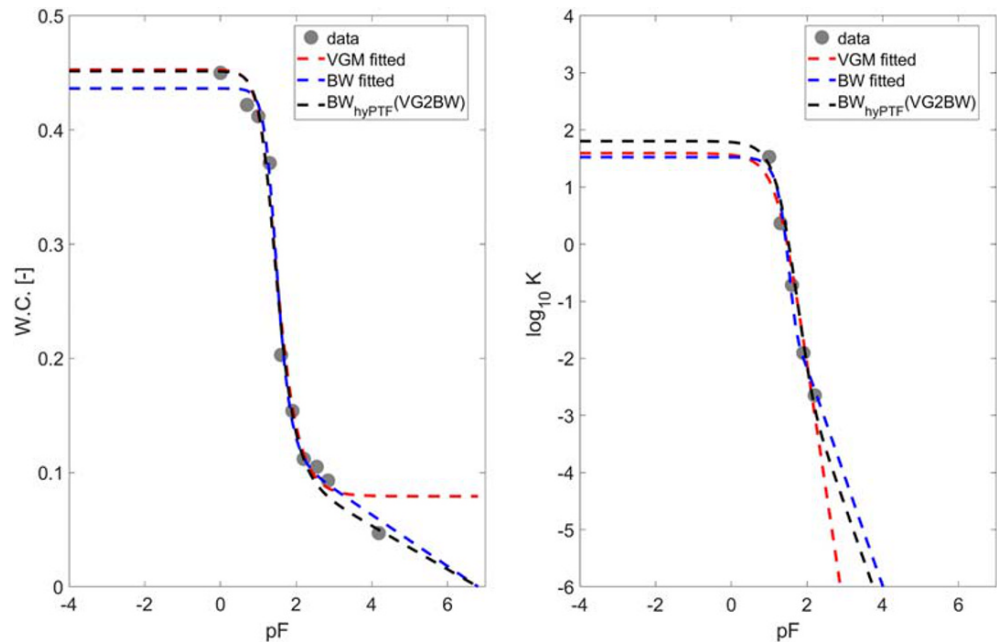
Figure 9. Exemplified comparative performance of the parameterized VGM model, the parameterized BW-VGM model, and the hyPTF-predicted BW model against the UNSODA 2573 data set. The remaining nine samples of the validation data set DS3 are shown in Figure A1.

by $\tau_{VGM} \ll 0$. Thus, we decided to take a functional approach, and use $\tau_{BW} = \min(\tau_{VGM}, 0)$; otherwise, the same approach is used as in Table 3.

3.3.2. K_{snc}

K_{snc} has no corresponding parameter in the VGM model. Initially, we tested the York regression (and several other methods, e.g., artificial neural networks and multiple linear regression, results not shown) to find a method to predict K_{snc} from VGM parameters. However, it was not possible to achieve a better predictive performance than with the median of $\widehat{K}_{snc}$, which we determined as $10^{-1.72}$ cm d^{-1}. Therefore, using the median appears to be the parsimonious choice. Diamantopoulos and Durner (2013) presented a model for corner flow based on theoretical considerations at the pore scale, whose maximum unsaturated conductivity value was determined as being very close to the value calculated here. Based on Ahe works by Tuller and Or (2001) and Diamantopoulos and Durner (2015) and the similarity in K_{snc} to the weighted saturated conductivity for the film flow component in Peters (2013), it appears not farfetched to construct informative priors in the near future. In particular, by considering the existing knowledge of the relationship between soil properties and saturated noncapillary flow, we are confident physical constraints can be incorporated. At this stage, we encourage the use of the published estimated hydraulic model parameters and their associated uncertainties in the accompanying data publication (cf. acknowledgement) in order to explore this further. Given current knowledge, we deem it premature to consider a prior in the Bayesian inference scheme for the parameter K_{snc} but encourage to use the presented data set to construct priors for future inference of parameters. It means that the value of $K_{snc} = 10^{-1.72}$ cm d^{-1} might change in the future and still needs further research.

3.3.3. Performance of the hyPTF

By way of example, the performance of the hyPTF is illustrated in Figure 9 for one of the samples (UNSODA 2573) of the randomly chosen *set-a-side* data set DS4. The results of the remaining nine samples from the validation data set can be inspected in Figure A1. Three different results of models are compared against the measured WRC and HCC data: (i) the VGM model as parameterized by inverse modeling, (ii) the same for the BW-VGM model, and (iii) the BW-VGM model parameterized by predicting the model parameters using the hyPTF established in the paper. Evidently, the inclusion of the noncapillary part in the WRC
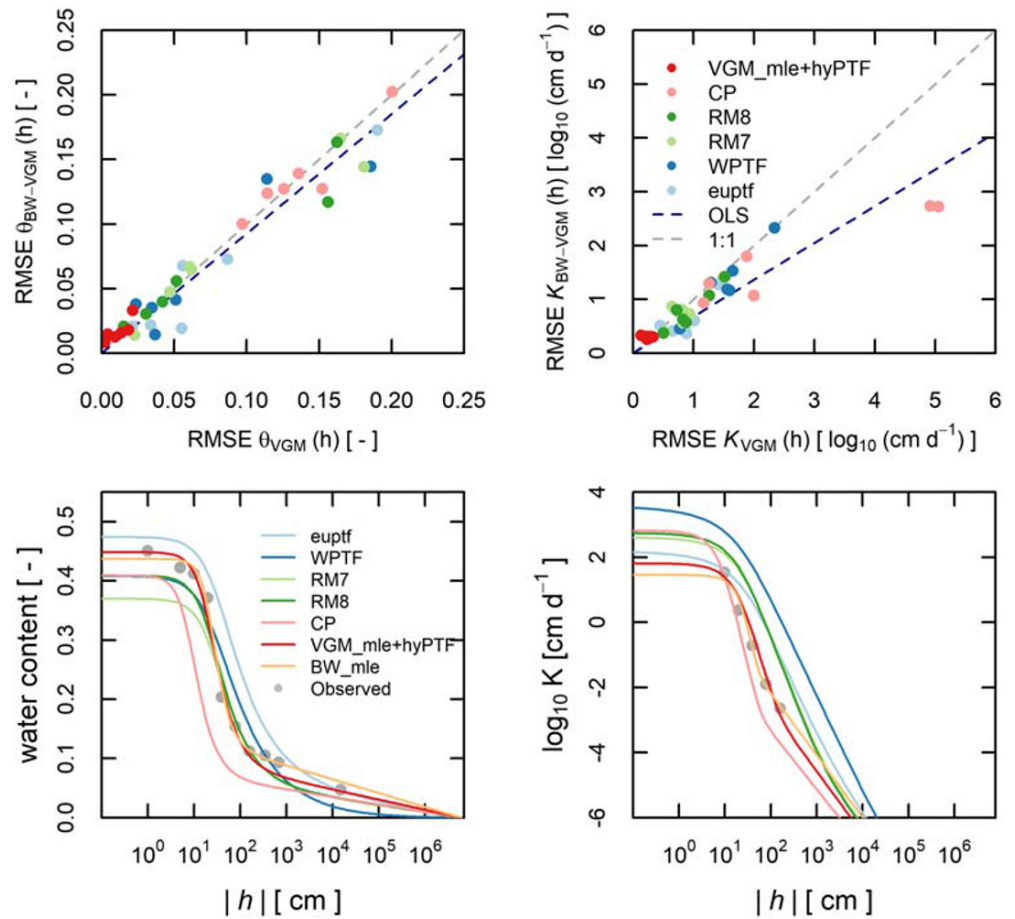
Figure 10. (Top row) Calculated RMSE for the volumetric water content and unsaturated hydraulic conductivity curves for the predicted VGM and the double-predicted hyPTF BW-VGM model for the samples from the validation data set with sufficient auxiliary information on soil properties for the use of the various literature hyPTF (abbreviations of models are explained in the text). (Bottom row) Performance of the double predicted BW-VGM model to describe the WRC (left) and HCC (data), exemplified on the sample UNSODA 2573 (further samples can be seen in Figure A2).

and HCC not only leads to the mentioned physically comprehensive model but also leads to a formidable description for $pF > 2$. This is also true for the BW model whose parameters were predicted applying knowledge of the estimated VGM parameters and the hyPTF.

To test the performance of the intended use of the hyPTF developed in this paper, we gathered five hyPTF from the literature. The five models were CP (Carsel & Parrish, 1988), RM7 (model H2m in Schaap et al., 2004), RM8 (H3m in Schaap et al., 2004), WPTF (Weynants et al., 2009 & Weihermüller et al., 2017), and euptf (Tóth et al., 2015) and made a double parameter prediction. In the initial step, we predicted the VGM model parameters for the six soils in DS4 which had soil property information. Based on those predicted parameters, and using the hyPTF from this paper, we predicted the BW-VGM parameters. The summary statistics are shown in Figure 10 (top row). The calculated RMSE results for both the WRC and the HCC between the modeled and measured data are displayed. In each, the RMSE value for the VGM model is given on the *x*-axis, and the RMSE of the corresponding BW-VGM model is given on the *y* axis. Visual inspection of Figure 10 clearly shows that the plotted RMSE fall below the 1:1 line for almost all cases in both data groups and, again, with a more pronounced effect for the HCC. This demonstrates that, in spite of a double prediction step, the BW-VGM model is still better than the VGM. For one sample, UNSODA 2573, we explicitly show the modeled WRC and HCC for (i) five double predicted parameter sets for the BW-VGM, (ii) the maximum likelihood estimate of the VGM-BW model

from maximizing Equation 12, and (iii) the hyPTF-predicted parameters of the BW-VGM model based on the maximum likelihood estimate of the VGM. We refrain from a discussion of the performance of the individual literature hyPTF at this stage. It follows that the RMSE shown in Figure 10 (top row) is within the range of other published PTFs (Nguyen et al., 2017; Román Dobarco et al., 2019; Zhang & Schaap, 2017).

We note that the presented double prediction hyPTF does not alleviate problems mentioned in the literature relating to unconsidered effects of stone content on SHPs (Nasri et al., 2015; Tetegan et al., 2015), on organic (Wallor et al., 2018), forest (Puhlmann & von Wilpert, 2012), or multimodal soils (Pollacco et al., 2017). In addition, the effects of dynamic nonequilibrium (Diamantopoulos et al., 2015; Šimůnek & van Genuchten, 2008) and macropore flow are not captured, since in the sense of the BW-VGM model, only an application to matrix flow is possible. Nevertheless, as the input hyPTF for the VGM model is region specific (e.g., Hollis et al., 2015; Román Dobarco et al., 2019), we are convinced our approach to work, here, too. Another advantage is that there are different VGM PTFs available (van Looy et al., 2017), with different requirements in input, so that, depending on the case specific information available, the case-specific double prediction step can be made. However, it is possible that this might not replace a newly constructed hyPTF based on machine learning tools, deemed superior for these kind of predictions (Szabó et al., 2019). A difficulty for that is that existing large data sets such as the EU-HYDI (Weynants et al., 2013) and HYPRESS (Wösten et al., 1999) data sets are not publically available, which in itself hampers model improvement.

4. Conclusion and Outlook

The aim of our study was to rigorously compare the VGM model with the physically comprehensive SHP model proposed by Weber et al. (2019). In their model framework, some important deficiencies in the VGM model are overcome, which is a result of a larger physical comprehensiveness. By using a large data set of 402 soils samples, it could be shown that this framework, the BW model, outperforms the VGM model in describing measured hydraulic properties. This is particularly visible in the HCC with a significant decrease in average MAE of the model-data matches. We could also demonstrate that it would be helpful for model development and assessment if WRC and HCC data were determined over a much wider range of pressure heads than those commonly included in available data sets.

Many environmental modeling codes, which adopt the Richards/Richardson equation to simulate water fluxes, are often implemented with the VGM and, also hyPTF, for the VGM. With the presented work, it will now be easy to use the better SHP models and the here presented new hyPTF. With this hyPTF, readily available data from soil maps (Dai et al., 2019) can be used to parameterize the BW model. The approach of the double prediction can also be used to permute input and algorithmic uncertainties (e.g., Carsel & Parrish, 1988; Deng et al., 2009; McNeill et al., 2018) through the hyPTF. In addition, Europe wide maps of BW-VGM model parameters based on available VGM maps (Szabó et al., 2019; Tóth et al., 2017) can easily be created, facilitating spatially explicit soil moisture modeling. This means that, for the first time, the application of a physically more comprehensive models may be applied for landscape-scale research.

In the worst case, the implementation of the BW model will lead to no significant improvement in simulated water redistribution in soils. In the best case, it can lead to more realistic water uplift from the groundwater or deeper soil zones. From this, we expect to have an effect on simulated plant growth, yield development and evapotranspiration when considering the soil-vegetation-atmosphere nexus. However, this requires a separate investigation. Finally, we are convinced that the presented hyPTF will (a) strengthen the utility of the BW model, (b) enable field-scale simulations of water and nutrient flux estimations in agroecosystems also under dry conditions, and (c) enable better impact assessment of warmer and drier meteorological conditions caused by climate change (at least for Western and Central Europe where these conditions have been predicted). In other words, it is possible that in modeling global food production, considerable differences may result from adopting a new model for the soil water fluxes. One step toward achieving this is by mapping model parameters based on soil property (Hengl et al., 2017) or hydraulic property (Tóth et al., 2017) maps.

APPENDIX
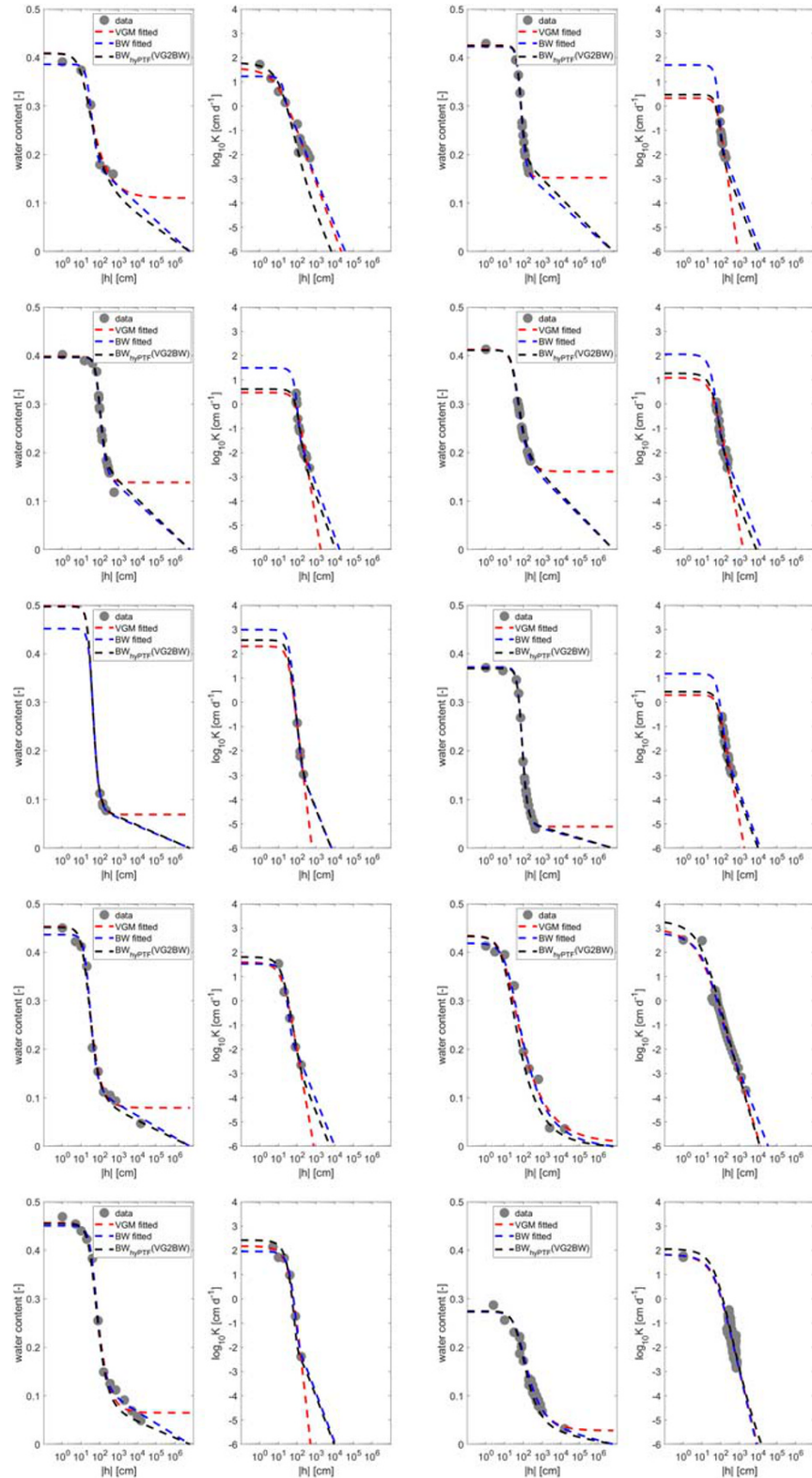
Figure A1. Like Figure 9, but for 10 samples. From top to bottom and left to right: UNSODA_4341, Schind_org_215, UNSODA_2216, UNSODA_2573, UNSODA_2763, Schind_org_604, Schind_min_631, Schind_org_371, UNSODA_4141, and Goncalves_6200006305.
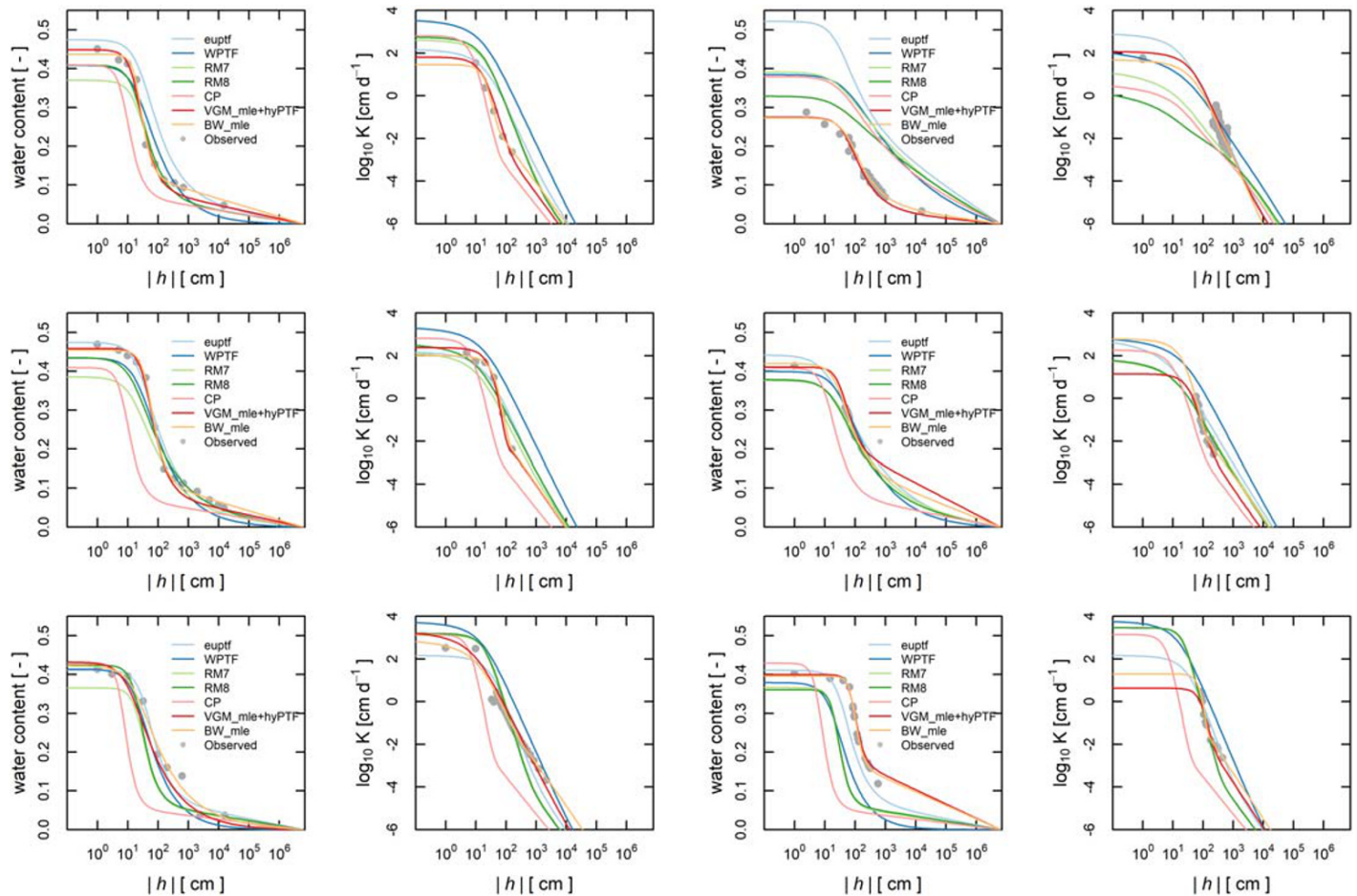
Figure A2. Results of the parameterized BW-VGM (BW), the hyPTF-predicted BW-VGM model parameters under knowledge of the parameterized VGM model for the respective sample, and the double-predicted parameters for the BW model from five different literature hyPTF (see details in text). The samples exhibited here are (top to bottom and left to right) UNSODA_2573, UNSODA_2763, UNSODA_4141, Goncalves_6200006305, Schind_min_631, and Schind_org_215.

Data Availability Statement

The data sets Weber, Finkel et al. (2020) used in this work can be freely accessed under http://hdl.handle.net/10900.1/c2f45822-86ba-4338-b3e0-2d068f024e07.

References

Abrahamsen, P., Hansen, S (2000). Daisy: An open soil-crop-atmosphere system model. *Environmental Modelling and Software*, *15*(3), 313–330. https://doi.org/10.1016/s1364-8152(00)00003-7

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723. https://doi.org/10.1109/TAC.1974.1100705

Angulo, C., Gaiser, T., Rötter, R. P., Børgesen, C. D., Hlavinka, P., Trnka, M., & Ewert, F. (2014). 'Fingerprints' of four crop models as affected by soil input data aggregation. *European Journal of Agronomy*, *61*, 35–48. https://doi.org/10.1016/j.eja.2014.07.005

Assouline, S., & Or, D. (2013). Conceptual and parametric representation of soil hydraulic properties: A review. *Vadose Zone Journal*, *12*. https://doi.org/10.2136/vzj2013.07.0121

Beyer, C., Altfelder, S., Duijnisveld, W. H. M., & Streck, T. (2009). Modelling spatial variability and uncertainty of cadmium leaching to groundwater in an urban region. *Journal of Hydrology*, *369*, 274–283. https://doi.org/10.1016/j.jhydrol.2009.02.022

Box, G. E. P., & Tiao, G. C. (1992). Nature of Bayesian inference. In G. E. P. Box, & G. C. Tiao (Eds.), *Bayesian inference in statistical analysis* (pp. 1–75). Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/9781118033197.ch1

Brooks, R. H. & Corey, A. T. (1964). Hydraulic properties of porous media, Hydrology Papers No 3, Colorado State University, Fort Collings, Colorado.

Brooks, S. P., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, *7*, 434–455. https://doi.org/10.1080/10618600.1998.10474787

Burnham, K. P., & Anderson, D. R. (2010). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed. p. 488). New York, NY: Springer.

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

Dai, Y., Shangguan, W., Wei, N., Xin, Q., Yuan, H., Zhang, S., et al. (2019). A review of the global soil property maps for Earth system models. *The Soil*, *5*, 137–158. https://doi.org/10.5194/soil-5-137-2019

Davison, A., & Hinkley, D. (1997). Bootstrap Methods and their Application (Cambridge Series in Statistical and Probabilistic Mathematics). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511802843

Deng, H., Ye, M., Schaap, M. G., & Khaleel, R. (2009). Quantification of uncertainty in pedotransfer function-based parameter estimation for unsaturated flow modeling. *Water Resources Research*, *45*, W04409. https://doi.org/10.1029/2008WR007477

Dettmann, U., & Bechtold, M. (2016). Deriving effective soil water retention characteristicsfrom shallow water table fluctuations in peatlands. *Vadose Zone Journal*, *15*. https://doi.org/10.2136/vzj2016.04.0029

Diamantopoulos, E., & Durner, W. (2012). Dynamic nonequilibrium of water flow in porous media: A review. *Vadose Zone Journal*, *11*. https://doi.org/10.2136/vzj2011.0197

Diamantopoulos, E., & Durner, W. (2013). Physically-based model of soil hydraulic properties accounting for variable contact angle and its effect on hysteresis. *Advances in Water Resources*, *59*, 169–180. https://doi.org/10.1016/j.advwatres.2013.06.005

Diamantopoulos, E., & Durner, W. (2015). Closed-form model for hydraulic properties based on angular pores with lognormal size distribution. *Vadose Zone Journal*, *14*. https://doi.org/10.2136/vzj2014.07.0096

Diamantopoulos, E., Durner, W., Iden, S. C., Weller, U., & Vogel, H.-J. (2015). Modeling dynamic non-equilibrium water flow observations under various boundary conditions. *Journal of Hydrology*, *529*, 1851–1858. https://doi.org/10.1016/j.jhydrol.2015.07.032

Diamantopoulos, E., Šimůnek, J., Oberdörster, C., Hammel, K., Jene, B., Schröder, T., & Harter, T. (2017). Assessing the potential exposure of groundwater to pesticides: A model comparison. *Vadose Zone Journal*, *16*. https://doi.org/10.2136/vzj2017.04.0070

Elmaloglou, S., Diamantopoulos, E., & Dercas, N. (2010). Comparing soil moisture under trickle irrigation modeled as a point and line source. *Agricultural Water Management*, *97*, 426–432. https://doi.org/10.1016/j.agwat.2009.11.003

Famiglietti, J. S., & Wood, E. F. (1994). Application of multiscale water and energy balance models on a tallgrass prairie. *Water Resources Research*, *30*, 3079–3093. https://doi.org/10.1029/94WR01499

Fatichi, S., Or, D., Walko, R., Vereecken, H., Young, M. H., Ghezzehei, T. A., et al. (2020). Soil structure is an important omission in Earth System Models. *Nature Communications*, *11*, 522. https://doi.org/10.1038/s41467-020-14411-z

Fredlund, D. G., & Xing, A. (1994). Equations for the soil-water characteristic curve. *Canadian Geotechnical Journal*, *31*, 521–532. https://doi.org/10.1139/t94-061

Gayler, S., Klier, C., Mueller, C. W., Weis, W., Winkler, J. B., & Priesack, E. (2009). Analysing the role of soil properties, initial biomass and ozone on observed plant growth variability in a lysimeter study. *Plant and Soil*, *323*, 125–141. https://doi.org/10.1007/s11104-009-9993-1

Gee, G. W., Campbell, M. D., Campbell, G. S., & Campbell, J. H. (1992). Rapid measurement of low soil water potentials using a water activity meter. *Soil Science Society of America Journal*, *56*, 1068–1070. http://www.scopus.com/inward/record.url?eid=2-s2.0-0027061667&partnerID=40&md5=8f878a13317552fa28ff5d26e9510594

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis, Texts in statistical science series* (Third ed.p. 667). Boca Raton, London, New York: CRC Press Taylor and Francis Group.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*, 457–472. https://doi.org/10.1214/ss/1177011136

Glaister, P. (2001). Least squares revisited. *The Mathematical Gazette*, *85*, 104. https://doi.org/10.2307/3620485

Goncalves, M. C., Leij, F. J., & Schaap, M. G. (2001). Pedotransfer functions for solute transport parameters of Portuguese soils. *European Journal of Soil Science*, *52*, 563–574. https://doi.org/10.1046/j.1365-2389.2001.00409.x

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*, 479–485. https://doi.org/10.1111/j.1365-2389.2004.00617.x

Haario, H., Laine, M., Mira, A., & Saksman, E. (2006). DRAM: Efficient adaptive MCMC. *Statistics and Computing*, *16*, 339–354. https://doi.org/10.1007/s11222-006-9438-0

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., et al. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE*, *12*, e0169748. https://doi.org/10.1371/journal.pone.0169748

Hoffmann, H., Zhao, G., Asseng, S., Bindi, M., Biernath, C., Constantin, J., et al. (2016). Impact of spatial soil and climate input data aggregation on regional yield simulations. *PLoS ONE*, *11*, e0151782. https://doi.org/10.1371/journal.pone.0151782

Höge, M., Wöhling, T., & Nowak, W. (2018). A primer for model selection: The decisive role of model complexity. *Water Resources Research*, *54*, 1688–1715. https://doi.org/10.1002/2017WR021902

Hollis, J. M., Lilly, A., Higgins, A., Jones, R. J. A., Keay, C. A., & Bellamy, P. (2015). Predicting the water retention characteristics of UK mineral soils. *European Journal of Soil Science*, *66*, 239–252. https://doi.org/10.1111/ejss.12186

Hopmans, J. W., Nielsen, D. R., & Bristow, K. L. (2002). How useful are small-scale soil hydraulic property measurements for large-scale vadose zone modeling? In P. A. C. Raats, D. Smiles, & A. W. Warrick (Eds.), *Environmental mechanics: Water, mass and energy transfer in the biosphere* (pp. 247–258). Washington, D. C: American Geophysical Union. https://doi.org/10.1029/129GM20

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*, 7530–7534. https://doi.org/10.1002/2014WR015937

Kashyap, R. L. (1982). Optimal choice of AR and MA parts in autoregressive moving average models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-4*, 99–104. https://doi.org/10.1109/TPAMI.1982.4767213

Kosugi, K. (1996). Lognormal distribution model for unsaturated soil hydraulic properties. *Water Resources Research*, *32*, 2697–2703. http://www.scopus.com/inward/record.url?eid=2-s2.0-0029663918&partnerID=40&md5=50b9b82551039064bfc9921b0991e9d2

Lawless, C., Semenov, M. A., & Jamieson, P. D. (2008). Quantifying the effect of uncertainty in soil moisture characteristics on plant growth using a crop simulation model. *Field Crops Research*, *106*, 138–147. https://doi.org/10.1016/j.fcr.2007.11.004

Martin, M. P., Lo Seen, D., Boulonne, L., Jolivet, C., Nair, K. M., Bourgeon, G., & Arrouays, D. (2009). Optimizing pedotransfer functions for estimating soil bulk density using boosted regression trees. *Soil Science Society of America Journal*, *73*, 485. https://doi.org/10.2136/sssaj2007.0241

McCarter, C. P. R., & Price, J. S. (2014). The hydrology of the Bois-des-Bel peatland restoration: Hydrophysical properties limiting connectivity between regenerated Sphagnum and remnant vacuum harvested peat deposit. *Ecohydrology*. https://doi.org/10.1002/eco.1498

McNeill, S. J., Lilburne, L. R., Carrick, S., Webb, T. H., & Cuthill, T. (2018). Pedotransfer functions for the soil water characteristics of New Zealand soils using S-map information. *Geoderma*, *326*, 96–110. https://doi.org/10.1016/j.geoderma.2018.04.011

Meskini-Vishkaee, F., Mohammadi, M. H., & Vanclooster, M. (2014). Predicting the soil moisture retention curve, from soil particle size distribution and bulk density data using a packing density scaling factor. *Hydrology and Earth System Sciences*, *18*, 4053–4063. https://doi.org/10.5194/hess-18-4053-2014

Millington, R. J., & Quirk, J. P. (1961). Permeability of porous solids. *Transactions of the Faraday Society*, *57*, 1200. https://doi.org/10.1039/TF9615701200

Mullen, K. M., Ardia, D., Gil, D. L., Windover, D., & Cline, J. (2011). DEoptim: An R package for global optimization by differential evolution. *Journal of Statistical Software*, *40*, 1–26. http://www.jstatsoft.org/v40/i06

Nasri, B., Fouché, O., & Torri, D. (2015). Coupling published pedotransfer functions for the estimation of bulk density and saturated hydraulic conductivity in stony soils. *Catena*, *131*, 99–108. https://doi.org/10.1016/j.catena.2015.03.018

Nemes, A., Schaap, M. G., Leij, F. J., & Wösten, J. H. M. (2001). Description of the unsaturated soil hydraulic database UNSODA version 2.0. *Journal of Hydrology*, *251*, 151–162. https://doi.org/10.1016/S0022-1694(01)00465-6

Nguyen, P. M., Haghverdi, A., de Pue, J., Botula, Y.-D., Le, K. V., Waegeman, W., & Cornelis, W. M. (2017). Comparison of statistical regression and data-mining techniques in estimating soil water retention of tropical delta soils. *Biosystems Engineering*, *153*, 12–27. https://doi.org/10.1016/j.biosystemseng.2016.10.013

Peters, A. (2013). Simple consistent models for water retention and hydraulic conductivity in the complete moisture range. *Water Resources Research*, *49*, 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*, 7535–7539. https://doi.org/10.1002/2014WR016107

Peters, A., Durner, W., & Wessolek, G. (2011). Consistent parameter constraints for soil hydraulic functions. *Advances in Water Resources*, *34*, 1352–1365. https://doi.org/10.1016/j.advwatres.2011.07.006

Philip, J. R. (1980). Field heterogeneity: Some basic issues. *Water Resources Research*, *16*, 443–448. https://doi.org/10.1029/WR016i002p00443

Pollacco, J. A. P., Webb, T., McNeill, S., Hu, W., Carrick, S., Hewitt, A., & Lilburne, L. (2017). Saturated hydraulic conductivity model computed from bimodal water retention curves for a range of New Zealand soils. *Hydrology and Earth System Sciences*, *21*, 2725–2737. https://doi.org/10.5194/hess-21-2725-2017

Priesack, E.: Expert-N dokumentation der modellbibliothek, Zugl.: Göttingen, Univ., Habil.-Schr., 2006, FAM-Bericht, 60, Hieronymus, München, pp. 298, 2006.

Puhlmann, H., & von Wilpert, K. (2012). Pedotransfer functions for water retention and unsaturated hydraulic conductivity of forest soils. *Journal of Plant Nutrition and Soil Science*, *175*, 221–235. https://doi.org/10.1002/jpln.201100139

Raats, P. A. C., & Knight, J. H. (2018). The Contributions of Lewis Fry Richardson to Drainage Theory, Soil Physics, and the Soil-Plant-Atmosphere Continuum. *Frontiers in Environmental Science*, *6*, 5. https://doi.org/10.3389/fenvs.2018.00013

Román Dobarco, M., Cousin, I., Le Bas, C., & Martin, M. P. (2019). Pedotransfer functions for predicting available water capacity in French soils, their applicability domain and associated uncertainty. *Geoderma*, *336*, 81–95. https://doi.org/10.1016/J.GEODERMA.2018.08.022

Romano, N., Hopmans, J., Simunek, J. J., & Durner, W. (2002). Simultaneous determination of water transmission and retention properties. *Inverse methods*, 963–1008.

Rudiyanto, S. M., van Genuchten, M. T., Alazba, A. A., Setiawan, B. I., & Minasny, B. (2015). A complete soil hydraulic model accounting for capillary and adsorptive water retention, capillary and film conductivity, and hysteresis. *Water Resources Research*, *51*, 8757–8772. https://doi.org/10.1002/2015WR017703

Saito, H., Šimůnek, J., & Mohanty, B. P. (2006). Numerical analysis of coupled water, vapor, and heat transport in the vadose zone. *Vadose Zone Journal*, *5*, 784. https://doi.org/10.2136/vzj2006.0007

Schaap, M. G., & Leij, F. J. (2000). Improved prediction of unsaturated hydraulic conductivity with the Mualem-van Genuchten model. *Soil Science Society of America Journal*, *64*, 843. https://doi.org/10.2136/sssaj2000.643843x

Schaap, M. G., Nemes, A., & van Genuchten, M. T. (2004). Comparison of models for indirect estimation of water retention and available water in surface soils. *Vadose Zone Journal*, *3*, 1455–1463. https://doi.org/10.2113/3.4.1455

Scharnagl, B., Vrugt, J. A., Vereecken, H., & Herbst, M. (2011). Inverse modelling of in situ soil water dynamics: Investigating the effect of different prior distributions of the soil hydraulic parameters. *Hydrology and Earth System Sciences*, *15*, 3043–3059. https://doi.org/10.5194/hess-15-3043-2011

Schelle, H., Heise, L., Jänicke, K., & Durner, W. (2013). Water retention characteristics of soils over the whole moisture range: A comparison of laboratory methods. *European Journal of Soil Science*, *64*, 814–821. https://doi.org/10.1111/ejss.12108

Schelle, H., Iden, S. C., & Durner, W. (2011). Combined transient method for determining soil hydraulic properties in a wide pressure head range. *Soil Science Society of America Journal*, *75*, 1681. https://doi.org/10.2136/sssaj2010.0374

Schneider, M., & Goss, K.-U. (2012). Prediction of water retention curves for dry soils from an established pedotransfer function: Evaluation of the Webb model. *Water Resources Research*, *48*, W06603. https://doi.org/10.1029/2011WR011049

Schöniger, A., Wöhling, T., Samaniego, L., & Nowak, W. (2014). Model selection on solid ground: Rigorous comparison of nine ways to evaluate Bayesian model evidence. *Water Resources Research*, *50*, 9484–9513. https://doi.org/10.1002/2014WR016062

Selim, H. M. E. (1998). *Physical nonequilibrium in soils: Modeling and application* (p. 492). Chelsea, Mich: Ann Arbor Press.

Šimůnek, J., & van Genuchten, M. T. (2008). Modeling Nonequilibrium Flow and Transport Processes Using HYDRUS. *Vadose Zone Journal*, *7*, 782. https://doi.org/10.2136/vzj2007.0074

Šimůnek, J., van Genuchten, M. T., & Šejna, M. (2016). Recent developments and applications of the HYDRUS computer software packages. *Vadose Zone Journal*, *15*. https://doi.org/10.2136/vzj2016.04.0033

Soetaert, K., & Petzoldt, T. (2010). Inverse modelling, sensitivity and Monte Carlo analysis in R using package FME. *Journal of Statistical Software*, *33*. https://doi.org/10.18637/jss.v033.i03

Storn, R., & Price, K. (1997). Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, *11*, 341–359. http://www.scopus.com/inward/record.url?eid=2-s2.0-0142000477&partnerID=40&md5=bdcec7e690cb1a1cbf3697cd70d7f28f

Streck, T., & Weber, T. K. D. (2020). Analytical expressions for noncapillary soil water retention based on popular capillary retention models. *Vadose Zone Journal*.

Szabó, B., Szatmári, G., Takács, K., Laborczi, A., Makó, A., Rajkai, K., & Pásztor, L. (2019). Mapping soil hydraulic properties using random-forest-based pedotransfer functions and geostatistics. *Hydrology and Earth System Sciences*, *23*, 2615–2635. https://doi.org/10.5194/hess-23-2615-2019

Tetegan, M., Richer de Forges, A. C., Verbeque, B., Nicoullaud, B., Desbourdes, C., Bouthier, A., et al. (2015). The effect of soil stoniness on the estimation of water retention properties of soils: A case study from central France. *Catena*, *129*, 95–102. https://doi.org/10.1016/j.catena.2015.03.008

Tokunaga, T. K. (2009). Hydraulic properties of adsorbed water films in unsaturated porous media. *Water Resources Research*, *45*, W06415. https://doi.org/10.1029/2009WR007734

Tóth, B., Weynants, M., Nemes, A., Makó, A., Bilas, G., & Tóth, G. (2015). New generation of hydraulic pedotransfer functions for Europe. *European Journal of Soil Science*, *66*, 226–238. https://doi.org/10.1111/ejss.12192

Tóth, B., Weynants, M., Pásztor, L., & Hengl, T. (2017). 3D soil hydraulic database of Europe at 250 m resolution. *Hydrological Processes*, *31*, 2662–2666. https://doi.org/10.1002/hyp.11203

Tuller, M., & Or, D. (2001). Hydraulic conductivity of variably saturated porous media: Film and corner flow in angular pore space. *Water Resources Research*, *37*, 1257–1276. https://doi.org/10.1029/2000WR900328

van Genuchten, M., Leij, F. J., & Wu, L. (1997) Characterization and measurement of the hydraulic properties of unsaturated porous media: Part 2.

van Genuchten, M. T. (1980). Closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, *44*, 892–898. http://www.scopus.com/inward/record.url?eid=2-s2.0-0019057216&partnerID=40&md5=9caf8cbebcd55775324a8ec4bd536e4a

van Genuchten, M. T., & Pachepsky, Y. A. (2011). Hydraulic properties of unsaturated soils. In J. Gliński, J. Horabik, & J. Lipiec (Eds.), *Encyclopedia of agrophysics* (pp. 368–376). Netherlands, Dordrecht: Springer. https://doi.org/10.1007/978-90-481-3585-1_69

van Looy, K., Bouma, J., Herbst, M., Koestel, J., Minasny, B., Mishra, U., et al. (2017). Pedotransfer functions in earth system science: Challenges and perspectives. *Reviews of Geophysics*, *55*, 1199–1256. https://doi.org/10.1002/2017RG000581

Vereecken, H., Maes, J., Feyen, J., & Darius, P. (1989). Estimating the soil moisture retention characteristics from texture, bulk density, and carbon content. *Soil Science*, *148*, 389–403. https://doi.org/10.1097/00010694-198912000-00001

Vereecken, H., Schnepf, A., Hopmans, J. W., Javaux, M., Or, D., Roose, T., et al. (2016). Modeling soil processes: Review, key challenges, and new perspectives. *Vadose Zone Journal*, *15*. https://doi.org/10.2136/vzj2015.09.0131

Vereecken, H., van Looy, K., Weynants, M., & Javaux, M. (2017). Soil retention and conductivity curve data base sDB, link to MATLAB files, supplement to: Weynants, Melanie; Vereecken, Harry; Javaux, Mathieu (2009): Revisiting vereecken pedotransfer functions: Introducing a closed-form hydraulic model. *Vadose Zone Journal*, *8*(1), 86–95.

Vereecken, H., Weihermüller, L., Assouline, S., Šimůnek, J., Verhoef, A., Herbst, M., et al. (2019). Infiltration from the pedon to global grid scales: An overview and outlook for land surface modeling. *Vadose Zone Journal*, *18*. https://doi.org/10.2136/vzj2018.10.0191

Vereecken, H., Weynants, M., Javaux, M., Pachepsky, Y., Schaap, M. G., & van Genuchten, M. T. (2010). Using pedotransfer functions to estimate the van Genuchten–Mualem soil hydraulic properties: A review. *Vadose Zone Journal*, *9*, 795. https://doi.org/10.2136/vzj2010.0045

Vogel, H.-J. (2019). Scale issues in soil hydrology. *Vadose Zone Journal*, *18*. https://doi.org/10.2136/vzj2019.01.0001

Vrugt, J. A. (2016). Markov Chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. *Environmental Modelling & Software*, *75*, 273–316. https://doi.org/10.1016/j.envsoft.2015.08.013

Wallor, E., Rosskopf, N., & Zeitz, J. (2018). Hydraulic properties of drained and cultivated fen soils part I—Horizon-based evaluation of van Genuchten parameters considering the state of moorsh-forming process. *Geoderma*, *313*, 69–81. https://doi.org/10.1016/j.geoderma.2017.10.026

Weber, T. K. D., Diamantopoulos, E., & Weynants, M. (2020). spsh: Estimation and prediction of parameters of various soil hydraulic property models: https://CRAN.R-project.org/package=spsh, R package version 1.1.0.

Weber, T. K. D., Durner, W., Streck, T., & Diamantopoulos, E. (2019). A modular framework for modelling unsaturated soil hydraulic properties over the full moisture range. *Water Resources Research*, *55*, 4994–5011. https://doi.org/10.1029/2018WR024584

Weber, T. K. D., Gerling, L., Reineke, D., Weber, S., Durner, W., & Iden, S. C. (2018). Robust inverse modeling of growing season net ecosystem exchange in a mountainous peatland: Influence of distributional assumptions on estimated parameters and total carbon fluxes. *Journal of Advances in Modeling Earth Systems*, *10*, 1319–1336. https://doi.org/10.1029/2017MS001044

Weber, T. K. D., Iden, S. C., & Durner, W. (2017a). Unsaturated hydraulic properties of Sphagnum moss and peat reveal trimodal pore-size distributions. *Water Resources Research*, *53*, 415–434. https://doi.org/10.1002/2016WR019707

Weber, T. K. D., Iden, S. C., & Durner, W. (2017b). A pore-size classification for peat bogs derived from unsaturated hydraulic properties. *Hydrology and Earth System Sciences*, *21*, 6185–6200. https://doi.org/10.5194/hess-21-6185-2017

Weber, T. K. D., Finkel, M., da Conceição Gonçalves, M., Vereecken, H., & Diamantopoulos, E. (2020). Measured soil hydraulic and physical properties and estimated soil hydraulic property model parameters. http://hdl.handle.net/10900.1/c2f45822-86ba-4338-b3e0-2d068f024e07

Weber, T. K. D., Vereecken, H., da Conceição Gonçalves, M., Diamantopoulos, E., & Finkel, M. (2020). Measured soil hydraulic properties: available after future acceptance, https://fdat.escience.uni-tuebingen.de/portal/

Weihermüller, L., Herbst, M., Javaux, M., & Weynants, M. (2017). Erratum to "Revisiting vereecken pedotransfer functions: Introducing a closed-form hydraulic model". *Vadose Zone Journal*, *16*. https://doi.org/10.2136/vzj2008.0062er

Weynants, M., Montanarella, L., Tóth, G., Arnoldussen, A., Anaya Romero, M., Bilas, G., et al. (2013). *European HYdropedological Data Inventory (EU-HYDI), EUR - Scientific and Technical Research series - ISSN* (pp. 1831-9424). Luxembourg.

Weynants, M., Vereecken, H., & Javaux, M. (2009). Revisiting vereecken pedotransfer functions: Introducing a closed-form hydraulic model. *Vadose Zone Journal*, *8*, 86. https://doi.org/10.2136/vzj2008.0062

Wösten, J. H. M., Lilly, A., Nemes, A., & Le Bas, C. (1999). Development and use of a database of hydraulic properties of European soils. *Geoderma*, *90*, 169–185. https://doi.org/10.1016/S0016-7061(98)00132-3

Ye, M., Meyer, P. D., & Neuman, S. P. (2008). On model selection criteria in multimodel analysis. *Water Resources Research*, *44*, W03428. https://doi.org/10.1029/2008WR006803

York, D. (1966). Least-squares fitting of a straight line. *Canadian Journal of Physics*, *44*, 1079–1086. https://doi.org/10.1139/p66-090

Zhang, Y., & Schaap, M. G. (2017). Weighted recalibration of the Rosetta pedotransfer model with improved estimates of hydraulic parameter distributions and summary statistics (Rosetta3). *Journal of Hydrology*, *547*, 39–53. https://doi.org/10.1016/j.jhydrol.2017.01.004