Journal of Hydrology 668 (2026) 135008

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

A differentiable hybrid modeling approach for learning soil water retention mechanisms from partial knowledge and data

Sarem Norouzi [a,c,*], Per Moldrup [b], Ben Moseley [c], David Robinson [d], Dani Or [e], Tobias L. Hohenbrink [f], Budiman Minasny [g], Morteza Sadeghi [h], Emmanuel Arthur [a], Markus Tuller [i,1], Mogens H. Greve [a], Lis W. de Jonge [a]

[a] Department of Agroecology, Aarhus University, Tjele, Denmark
[b] Department of the Built Environment, Aalborg University, Aalborg, Denmark
[c] Department of Earth Sciences and Engineering, Imperial College London, London, UK
[d] UK Centre for Ecology & Hydrology, Bangor, UK
[e] Department of Civil and Environmental Engineering, University of Nevada, Reno, NV, USA
[f] German Weather Service (DWD), Agrometeorological Research Center, Braunschweig, Germany
[g] School of Life and Environmental Sciences, The University of Sydney, Australia
[h] California Department of Water Resources, Sacramento, USA
[i] Department of Environmental Science, The University of Arizona, Tucson, AZ, USA

ARTICLE INFO

This manuscript was handled by Renato Morbidelli, Editor-in-Chief

Keywords:
Scientific machine learning
Differentiable modeling
Soil hydraulic properties
Hybrid modeling

ABSTRACT

Soil physics models have long relied on simplifying assumptions to represent complex processes, yet such assumptions can strongly bias model predictions. Here, we propose differentiable hybrid modeling (DHM) as a paradigm-shifting framework that learns unobservable intrinsic processes from data and physical constraints, rather than simplifying them. As a proof of concept, we apply the DHM approach to the challenge of partitioning the soil water retention curve (SWRC) into capillary and adsorbed water components, a problem where traditional assumptions have led to divergent results. The hybrid framework derives this partitioning directly from data while remaining guided by simple physical constraints. Using basic soil physical properties as inputs, the DHM couples an analytical formula for the dry end of the SWRC with data-driven physics-informed neural networks that learn the wet end, the transition between the two ends, and key soil-specific parameters. The model was trained on a SWRC dataset from 482 undisturbed soil samples, spanning a broad range of texture classes and organic carbon contents. The hybrid model successfully learned both the overall shape and the capillary and adsorbed components of the SWRC. Notably, the learned patterns were consistent with pore-scale thermodynamic saturation behavior in angular pores, without relying on explicit assumptions about soil pore geometry or its distribution. Moreover, the model revealed a distinctly nonlinear transition between capillary and adsorbed domains, challenging the linear assumptions invoked in previous studies. The methodology introduced here provides a blueprint for learning other soil processes where high-quality datasets are available but mechanistic understanding is incomplete.

1. Introduction

Physics-based modeling approaches have a long tradition in soil physics and have been applied to simulate fundamental soil processes such as infiltration, evaporation, solute transport, and energy exchange in the vadose zone (Green and Ampt, 1911; Gardner, 1958; Philip and de De Vries, 1957; Van Genuchten, 1982). Mechanistic modeling typically involves conceptualizing the problem, deriving governing equations from physical laws or empirical relationships, and validating the resulting models against experimental data. These models are then used to study system functions and behaviors, test hypotheses, and assess the responses of a system to changes in the driving forces or internal properties.

The process of deriving representative models in any natural system inevitably requires simplifying poorly understood components of the

* Corresponding author at: Department of Agroecology, Aarhus University, Tjele, Denmark.
E-mail address: sarem.nrz@agro.au.dk (S. Norouzi).
[1] The late Markus Tuller contributed substantially to this research; he passed away before the manuscript was submitted.

https://doi.org/10.1016/j.jhydrol.2026.135008
Received 30 October 2025; Received in revised form 19 January 2026; Accepted 20 January 2026
Available online 21 January 2026

Nomenclature

Notation	Definition Unit
θ	Total volumetric water content [$m^3\ m^{-3}$]
θ_c	Volumetric capillary water content (capillary component) [$m^3\ m^{-3}$]
θ_a	Volumetric adsorbed film water content (adsorbed component)]$m^3\ m^{-3}$]
h	Matric head [cm]
pF	Log-transformed matric head [-]
θ_{CS}	Water content predicted by the Campbell–Shiozawa model [$m^3\ m^{-3}$]
pF_{dry}	pF at oven dryness where $\theta \to 0$ [-]
θ_o	Campbell–Shiozawa fitting parameter [$m^3\ m^{-3}$]
θ_s	Volumetric saturated water content [$m^3\ m^{-3}$]
S_c	Capillary saturation, $S_c = \theta_c/\theta_s$ [-]
$f(S_c)$	Transition function [-]
$g(S_c)$	Unknown function for parametrizing the transition function [-]
$\mathbf{x}$	Input vector of basic soil properties [$Sand, Silt, Clay, OC, BD$] [see text]
NN_c, NN_g, NN_s, NN_{dry}, NN_o	Neural sub-models for predicting θ_c, $g(S_c)$, θ_s, pF_{dry}, θ_o [see text]
$\boldsymbol{\phi}_c$, $\boldsymbol{\phi}_g$, $\boldsymbol{\phi}_s$, $\boldsymbol{\phi}_{dry}$, $\boldsymbol{\phi}_o$	Trainable parameters of NN_c, NN_g, NN_s, NN_{dry}, NN_o [-]
J	Loss function [-]
$\widehat{\theta}$	Predicted volumetric water content [$m^3\ m^{-3}$]
N	Total number of measured points [-]
N_{wet}, N_{dry}	Number of wet-end and dry-end training examples [-]
λ_1, λ_2, λ_3, λ_4	Loss-term weights [-]
S_1, S_2	Number of residual (collocation) points [-]
$pF_{air-entry}$	Air-entry value [-]
$\overline{\theta}$	Mean of measured water contents [$m^3\ m^{-3}$]
Q_{25}, Q_{75}	25th and 75th percentiles of measured θ [$m^3\ m^{-3}$]
θ_{co}	Water content at the crossover point between capillary and adsorbed components [$m^3\ m^{-3}$]
ρ_b	Soil bulk density [$g\ cm^{-3}$]

system. For instance, soil physics makes assumptions regarding (difficult to quantify) pore geometry and its distribution within soil, attainment of equilibrium conditions, or the functional forms of constitutive relationships. While these simplifications enable tractable formulations, they are shaped by the modeler's view of the system that may bias the true representation of soil processes. While some assumptions are refined as new evidence emerges, the original modeling and representation bias may persist.

To overcome some of these challenges, we propose differentiable hybrid modeling (DHM) as an alternative approach to the explicit representation of certain soil physical processes. Hybrid methods embed neural networks within physical models so that the unknown or poorly understood components of a system can be learned directly from data, while the well-established physical laws remain explicitly enforced (Psichogios and Ungar, 1992; Karniadakis et al., 2021; Moseley, 2022; Shen et al., 2023). By constructing both the physical equations and the neural components in a differentiable form, these hybrid systems can be trained end-to-end using gradient-based optimization. This allows all parameters to be adjusted jointly to minimize a downstream, physics-constrained loss function. This dual nature of these methods maintains the interpretability of traditional formulations while enabling discovery of processes that are otherwise inaccessible through purely mechanistic or purely empirical approaches.

Automatic differentiation (AD), which is the backbone of DHM (Baydin et al., 2018), has also advanced other domains of scientific machine learning (SciML), such as physics-informed neural networks (PINNs) (Raissi et al., 2019). PINNs have been successfully applied in vadose zone modeling to estimate soil hydraulic properties from soil moisture measurements (Tartakovsky et al., 2020; Bandai and Ghezzehei, 2021; Minasny et al., 2024), to model water flow and solute transport using geoelectrical data (Haruzi and Moreno, 2023), and to develop flexible non-parametric pedotransfer functions (PTFs) for the soil water retention curve (Norouzi et al., 2025). While both DHM and PINNs fall under the broader umbrella of SciML and share conceptual similarities, they differ fundamentally in structure and training objectives. In PINNs, neural networks are trained to learn solutions that satisfy governing physical laws, whereas DHM embeds neural networks within an existing analytical model or process-based model to learn the unknown or less understood components.

To demonstrate how DHM can be used as a framework for learning complex soil physical processes, we apply this method to the classic problem of modeling the soil water retention curve and its partitioning to capillary and adsorptive components. The SWRC is a fundamental soil property that defines the relationship between water content and matric head (or suction) in soil. Modeling a variety of land-surface processes, including infiltration, runoff, evaporation, and energy exchange at the land surface requires knowledge of the SWRC across scales (Gupta et al., 2022; Tehrani et al., 2025; Turek et al., 2025). Moreover, partitioning the SWRC into its capillary and adsorbed components enhances the modeling of the soil hydraulic conductivity curve and is essential for determining the liquid–water interfacial area, which influences soil health and biogeochemical processes, particularly the retention and transport of interfacially active contaminants such as per- and polyfluoroalkyl substances (PFAS) (Guo et al., 2020; Brusseau, 2023).

All existing SWRC models that partition capillary and adsorbed film water have been developed based on specific prior assumptions and simplifications about soil pore geometry, the functional forms used to describe water retention components, and the transition between these two soil water regimes. Because these assumptions directly influence the resulting partitioning, different models applied to the same soil can yield substantially different outcomes, reflecting the sensitivity of predictions to their underlying assumptions (Or and Tuller, 1999; Lebeau and Konrad, 2010; Peters, 2013; Lu, 2016; Weber et al., 2019; Ghorbani et al., 2025). Moreover, direct experimental validation of the partitioning remains challenging, as capillary and adsorbed water components are not readily separable or independently measurable at the pore scale. Therefore, there remains a need for new approaches that can learn the partitioning directly from data with minimal prior assumptions, while still respecting key physical constraints.

The main objective of this study is to develop and demonstrate a hybrid modeling framework that learns the shape of the SWRC as well as its capillary and adsorbed water components from basic soil properties using a state-of-the-art differentiable modeling approach. In key contrast to traditional parametric models with rigid physical assumptions, our hybrid method learns a flexible, physically interpretable, and data-driven partitioning, with only universally accepted assumptions included. Our study introduces a new generation of SWRC models, which we term "semi-parametric" (i.e., semi-analytical) models. In this new type of SWRC models, part of the curve is described by analytical equations, while the remaining parts are flexibly learned from data using neural networks, yet the final SWRC remains continuous, differentiable, and physically consistent. We believe our work demonstrates the potential of the DHM framework for unifying physical theory and data-driven discovery across a wide range of fundamental soil physical processes.

2. Materials and methods

2.1. Hybrid modeling framework

The hybrid model uses the same inputs and outputs as conventional pedotransfer functions, translating basic soil physical properties to SWRC. However, during this mapping from inputs to outputs, the model learns multiple intermediate processes and their unobservable transitions implicitly, without requiring explicit data for them.

We begin with universally accepted principles describing the coexistence and behavior of capillary and adsorptive forces across the full moisture range. As we derive the governing equations from these known principles, any unknown terms are treated as components to be inferred from data using physics-constrained neural networks. To physically guide the model, we only rely on a few physical constraints that are universally accepted and relax the restrictive assumptions in conventional models: idealized pore shapes, certain pore size distributions, or fixed functional forms.

2.1.1. Physical definitions and model derivation

Capillary water refers to liquid water filling the spaces between soil particles, held by surface tension and the contact angle of water with solid surfaces, which leads to the formation of curved liquid–vapor interfaces (menisci). The adsorbed film water component refers specifically to liquid water retained in thin films by adsorptive forces, where a distinct liquid–air interface is present (see for example figure 4 of Nachum (2025)). The adsorptive forces in soil arise from intermolecular interactions between the liquid and solid surfaces, including van der Waals forces, electrostatic double-layer forces, and structural (hydration) forces (Derjaguin et al., 1987; Tuller and Or, 2005a).

The total volumetric water content retained in a soil can thus be expressed as the sum of its capillary (θ_c) and adsorbed (θ_a) components:

$$\theta = \theta_c + \theta_a \tag{1}$$

Note that all terms in Eq. (1) are functions of matric head. As a soil dries, capillary water drains from larger pores and recedes into pore corners. Beyond a certain matric head threshold, water persists primarily as thin films adsorbed onto particle surfaces. At this dry end, experimental studies show that the water retention curve becomes linear in $pF-\theta$ space, where $pF = \log|h|$ and h is the matric head in cm (Arthur et al., 2013). This linear behavior can be described analytically by the Campbell and Shiozawa (1992) model (hereinafter referred to as Campbell-Shiozawa model and denoted by the subscript CS), which in $pF-\theta$ space is written as:

$$\theta_{CS} = \left(1 - \frac{pF}{pF_{dry}}\right)\theta_o \tag{2}$$

where θ_{CS} is the predicted water content by Campbell-Shiozawa model, and θ_o and pF_{dry} are its fitting parameters that determine the slope and intercept of this empirical model at the dry end. The parameter pF_{dry} corresponds to the logarithm of matric head at oven dryness, where the soil is assumed to reach zero water content.

The Campbell–Shiozawa model was originally developed for the dry end of the SWRC (i.e., the higher range of pF values), where only adsorptive forces are active. In the lower range of pF values, where capillary water begins to contribute, the expression for θ_{CS} no longer holds and must be revised. To account for this, we introduce a transition function, denoted as f, that modifies θ_{CS} in the mixed region, where capillary and adsorbed water coexist. This function is treated as an unknown to be learned from data, and it is expressed as a function of capillary saturation, defined as $S_c = \theta_c/\theta_s$, where θ_s is the saturated water content. Accordingly, the adsorbed film component of the SWRC can be modeled as:

$$\theta_a = f(S_c)\theta_{CS} \tag{3}$$

Table 1
Summary of the architecture and configuration of the subnetworks.

Network	Inputs[a]	Hidden layers[b]	Output	Output activation and scaling[c]
NN_c	*Sa*, *Si*, *Cl*, *OC*, *BD*,*pF*	2 × Dense (8) (ELU)	θ_c	$\theta_s\sigma(z_c)$
NN_g	NN_c/NN_s	2 × Dense (16) (ELU)	$g(S_c)$	Linear
NN_s	*Sa*, *Si*, *Cl*, *OC*, *BD*	2 × Dense (4) (ELU)	θ_s	$\sigma(z_s)$
NN_o	*Sa*, *Si*, *Cl*, *OC*, *BD*	2 × Dense (4) (ELU)	θ_o	$\sigma(z_o)$
NN_{dry}	*Sa*, *Si*, *Cl*, *OC*, *BD*	2 × Dense (4) (ELU)	pF_{dry}	$6.2 + \sigma(z_{dry}) \times (7.6 - 6.2)$

[a] *Sa*, *Si*, *Cl*, *OC*, and *BD* stand for sand, silt, clay, organic carbon, and bulk density, respectively.

[b] This column represents the number of hidden layers, the layer type, the number of units per layer, and the activation function for each neural network (*NN*).

[c] $\sigma(z_x)$ represents the sigmoid activation function applied to the raw output z_x of the neural network associated with variable x.

Combining Eqs. (3) and (1), the total water can be expressed as:

$$\theta = \theta_c + \overbrace{f(S_c)\theta_{CS}}^{\theta_a} \tag{4}$$

The second term on the right-hand side of Eq. (4) accounts for the adsorbed film water contribution. When S_c approaches zero (i.e., at very high pF values where θ_c is zero), the retention behavior is dominated by adsorbed water. In this limit, Eq. (4) should reduce to the Campbell-Shiozawa model, Eq. (2). To ensure this, we reparametrize the transition function with a hard constraint that enforces $f(0) = 1$:

$$f(S_c) = 1 + S_c \bullet g(S_c) \tag{5}$$

where $g(S_c)$ is an unknown function that is learned from data. Combining this equation with Eq. (4), the total water content can be expressed as:

$$\theta = \theta_c + \overbrace{[1 + S_c g(S_c)]\theta_{CS}}^{\theta_a} \tag{6}$$

Inserting Eq. (2) into Eq. (6), we obtain:

$$\theta = \theta_c + [1 + S_c g(S_c)]\left(1 - \frac{pF}{pF_{dry}}\right)\theta_o \tag{7}$$

It is worth noting that Eq. (7) reflects only a general structural formulation based on a few widely accepted assumptions about water retention in soils. Specifically, it assumes that total water content consists of two components, capillary and adsorbed water, and that the contribution of the capillary component vanishes as capillary saturation (S_c) approaches zero. In this condition, Eq. (7) reduces to Campbell-Shiozawa model for the dry end. The function $g(S_c)$ captures the transition between capillary and adsorbed dominant regions behavior without explicitly specifying its shape and form in advance. Similarly, the capillary component, θ_c, as well as soil constants (i.e., θ_s, pF_{dry}, and θ_o) are treated as unknowns to be learned from data.

2.1.2. Neural sub-model for the capillary water

We replace the capillary water content (θ_c) in Eq. (7), which is a function of pF, with a dedicated neural network sub-model named $NN_c(\mathbf{x}, pF; \boldsymbol{\phi}_c)$, which receives the vector of soil physical properties $\mathbf{x} = [Sand, Silt, Clay, OC, BD]$ and pF as input. The vector $\boldsymbol{\phi}_c$ represents the trainable parameters (weights and biases) of this neural network. This network has two hidden layers, each with eight units (i.e., nodes) (Table 1). By taking pF as an input, the network outputs the capillary

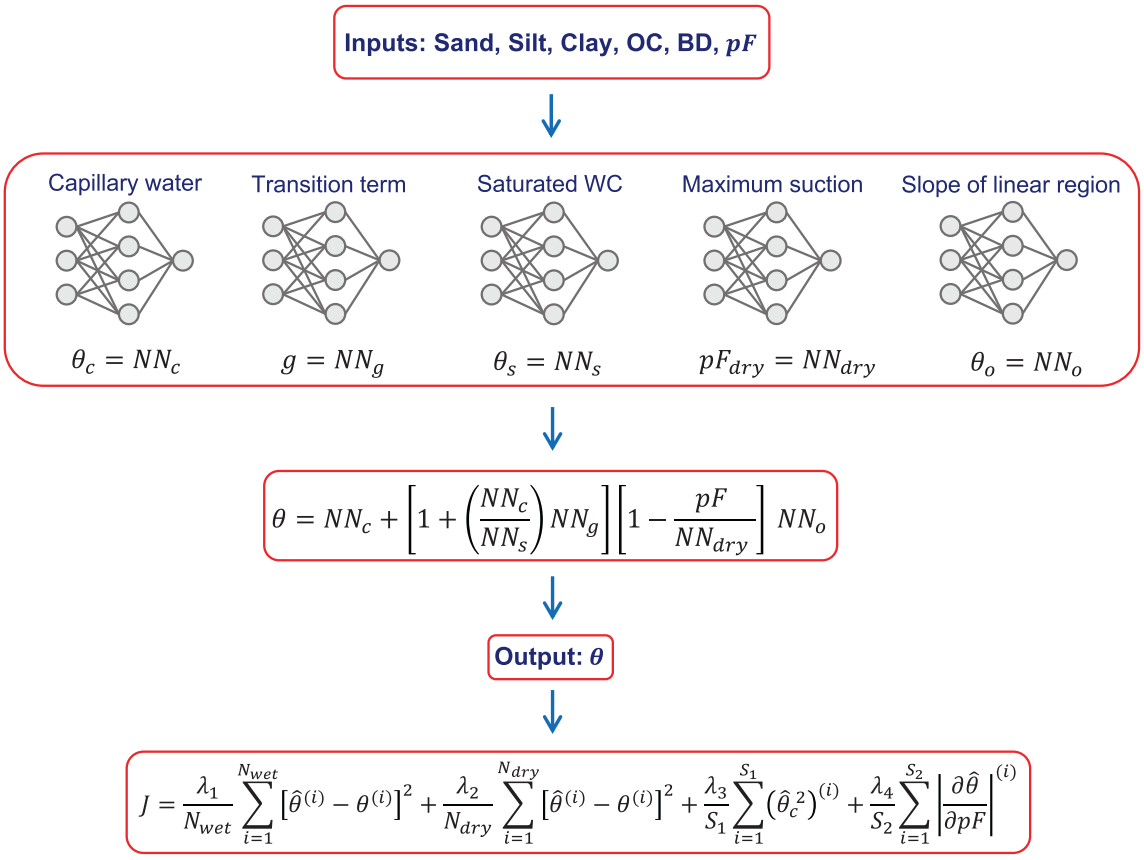

Fig. 1. Workflow of the proposed hybrid model illustrating the interior neural networks, analytical formulation, and loss function structure for learning the soil water retention curve and its capillary and adsorbed components. The model inputs, including soil physical properties and *pF*, are fed into separate neural networks whose outputs are combined within the hybrid formulation to predict the total water content, which is then used in the loss function. The entire model, including all neural network components, is trained end-to-end, meaning that all parameters are optimized jointly from input to output through gradient-based minimization of a downstream, physics-constrained loss function, without explicit labeled data for the individual sub-networks. *pF* serves as an input only to the capillary neural network (NN_c).

water content at the specified *pF*. This architecture produces a continuous representation of the curve without limiting it to a specific functional form (Haghverdi et al., 2012; Norouzi et al., 2025). To ensure meaningful predictions, the sub-network output is constrained to remain less than saturated water content (θ_s). To enforce this, the raw output of the capillary network is passed through a sigmoid activation, $\sigma(z) = 1/[1 + \exp(-z)]$, and scaled by the θ_s value.

2.1.3. Neural sub-model for the transition function

The transition function in Eqs. (5) and (7) includes an unknown component, $g(S_c)$, which maps the input S_c to a scalar output. Rather than assuming a fixed analytical form for $g(S_c)$, we replace it with a fully connected neural network, $NN_g\left(S_c;\boldsymbol{\phi}_g\right)$, and allow it to be learned from data. Similarly, $\boldsymbol{\phi}_g$ refers to the trainable parameters of this network. This neural network is capable of approximating a wide range of continuous functions, which helps capture complex transition behaviors. The neural network $NN_g\left(S_c;\boldsymbol{\phi}_g\right)$ consists of two hidden layers and a linear output layer without any constraint on the output value (Table 1). We designed this sub-network to be more flexible to ensure that the shape of the overall transition function, $f(S_c)$, is not limited by the capacity of $NN_g\left(S_c;\boldsymbol{\phi}_g\right)$.

2.1.4. Neural sub-model for the soil dependent constants

The Campbell–Shiozawa model and the transition function both depend on three key soil-specific parameters: θ_s, pF_{dry}, and θ_o. We assume that these parameters can be predicted from basic soil physical properties, and therefore, we model each parameter using a dedicated neural network. In our hybrid framework, each parameter is estimated by a separate sub-network: $NN_s(\mathbf{x};\boldsymbol{\phi}_s)$, $NN_{dry}\left(\mathbf{x};\boldsymbol{\phi}_{dry}\right)$, and $NN_o(\mathbf{x};\boldsymbol{\phi}_o)$, corresponding to θ_s, pF_{dry}, and θ_o, respectively. Each sub-network takes five basic soil properties as defined by $\mathbf{x}$: sand, silt, clay, OC, and BD. Additionally, each of the neural networks has two hidden layers, each containing four units with Exponential Linear Unit (ELU) activation function (Table 1). Although a single multi-output neural network could have been used, given that the inputs of these networks are identical, we opted for separate networks to maintain clarity and interpretability in the modeling framework.

Since θ_s and θ_o represent volumetric water contents and cannot exceed unity, we constrain the predictions of $NN_s(\mathbf{x};\boldsymbol{\phi}_s)$ and $NN_o(\mathbf{x};\boldsymbol{\phi}_o)$ to a range between 0 and 1 by applying a sigmoid activation function in the output layers of these sub-networks. Similarly, several studies showed that the range of pF_{dry} falls between 6.5 and 7.45 (Schneider and Goss, 2012; Arthur et al., 2013; Lu and Khorshidi, 2015; Karup et al., 2017). To ensure broader applicability and account for potential variability beyond these observations, we constrain the $NN_{dry}\left(\mathbf{x};\boldsymbol{\phi}_{dry}\right)$ sub-model output to predict values within a slightly wider interval of 6.2 to 7.6. This is achieved by scaling the sigmoid output of $NN_{dry}\left(\mathbf{x};\boldsymbol{\phi}_{dry}\right)$ to this target range.

To keep notation simple, the explicit dependence on inputs $(\mathbf{x}, pF)$ and trainable parameters $(\boldsymbol{\phi})$ is omitted in the remaining text wherever it does not cause ambiguity.

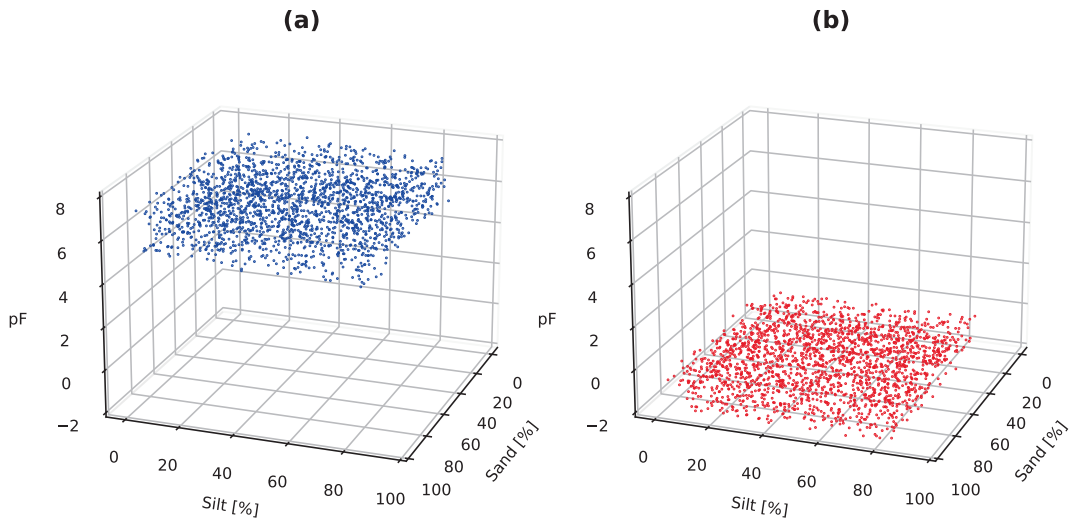

Fig. 2. The two sets of residual (collocation) points used to enforce physical constraints: (a) zero capillary water for $pF > 5$ and (b) constant water content for $pF < -0.3$.

2.1.5. Hybrid model and loss function

By embedding all the neural network sub-models into Eq. (7), we obtain our hybrid model, expressed as:

$$\theta = NN_c + \overbrace{\left[1 + \left(\frac{NN_c}{NN_s}\right) NN_g\right]\left(1 - \frac{pF}{NN_{dry}}\right) NN_o}^{\theta_a} \tag{8}$$

As depicted in Fig. 1, the model inputs, including soil physical properties and pF, are fed into separate neural networks whose outputs are combined within the hybrid formulation to predict the total water content, which is then used in the loss function.

It is important to note that Eq. (8) is directly trained end-to-end on measurements of the total water content as a function of pF (Fig. 1), and all the neural network parameters are learned jointly. Therefore, separate training of individual sub-networks using distinct target measurements is not needed. Instead, the sub-networks adjust their predictions during training so that the final predicted total water content closely matches the measured total water content at each pF value (Fig. 1).

In our hybrid modeling design, each measured point on the SWRC is treated as an independent training example, paired with its corresponding soil properties: sand, silt, clay, OC, BD, and pF. The key assumption is that, given a sufficient number of measurements, training the hybrid model in Eq. (8) enables the model to capture the complete shape of the SWRC, including both the capillary and adsorbed water components, which are modeled with various internal sub-networks. A major advantage of this method is that it allows samples with very few measured points to be included in the training set. This capability is not achievable with conventional parametric models, where a minimum number of measured points from each sample is needed (Rasoulzadeh et al., 2025).

The loss function used for training the hybrid model is as follows:

$$J = \frac{\lambda_1}{N_{wet}} \sum_{i=1}^{N_{wet}} \left[\widehat{\theta}^{(i)} - \theta^{(i)}\right]^2 + \frac{\lambda_2}{N_{dry}} \sum_{i=1}^{N_{dry}} \left[\widehat{\theta}^{(i)} - \theta^{(i)}\right]^2 + \frac{\lambda_3}{S_1} \sum_{i=1}^{S_1} \left(\widehat{\theta}_c^2\right)^{(i)} + \frac{\lambda_4}{S_2} \sum_{i=1}^{S_2} \left|\frac{\partial \widehat{\theta}}{\partial pF}\right|^{(i)} \tag{9}$$

where $\widehat{\theta}$ and θ are the predicted and measured water contents, respectively, and they are both a function of pF. The first two terms on the right-hand side of Eq. (9) represent the mean squared error between the volumetric water contents predicted by the neural network and the observed measurements, where N_{wet} and N_{dry} denote the number of training examples from the wet and dry ends, respectively. As shown by Norouzi et al. (2025), using separate terms for the wet-end (i.e., $pF \leq 4.2$) and dry-end (i.e., $pF > 4.2$) is necessary to account for disparities in sample sizes and the narrower range of water contents typically observed at the dry end. The λ coefficients are weights assigned to each term in the loss function which determine the relative importance of each loss component during training. The parameters S_1 and S_2 in the last two terms of Eq. (9) denote the number of residual (collocation) points used to enforce the two physics-based constraints, which are explained in the following section.

2.1.6. Universally accepted physical constraints

Our hybrid model relies solely on general physical reasoning without imposing rigid or system-specific assumptions. Here, to guide the model, we incorporate two physical constraints that are broadly accepted in soil physics and supported by pore-scale saturation mechanisms.

First, at high suctions in soil (i.e., for $pF > 5$), the soil water content is assumed to be entirely in adsorbed form, meaning that the capillary water content should approach zero in this range of pF (Norouzi et al., 2025; Tuller and Or, 2005b). To enforce this, we introduce a constraint in the loss function. Specifically, we generate a set of residual (collocation) points, which are synthetic samples with random combinations of sand, silt, clay, OC, and BD, paired with random pF values higher than 5. During each training step, the output of the θ_c neural network at these residual points is computed, and the mean of the squared values is added as a penalty term to the custom loss function to encourage θ_c to approach zero for $pF > 5$ (the third term on the right-hand side of Eq. (9). These residual points are illustrated in Fig. 2a.

Second, the soil air-entry value, also known as the bubbling pressure, corresponds to the matric head at which air begins to penetrate the largest soil pores (Fredlund and Xing, 1994; Sourmanabad et al., 2024). According to this definition, when the soil matric head (expressed in terms of pF) is below a specific value, the soil remains saturated, and its water content remains constant despite further changes in matric head. This condition can be mathematically represented as:

$$\frac{d\theta}{dpF} = 0, pF < pF_{air-entry} \tag{10}$$

The air-entry value, $pF_{air-entry}$, depends on both soil texture and structure, with structure often playing a dominant role in undisturbed samples (Rawls et al., 1982). To implement this constraint within our neural network, we set $pF = -0.3$ (equivalent to a matric head of -0.5 cm) as the minimum threshold, below which the soil is assumed to stay fully saturated, with water content remaining constant despite variations in pF. It should be noted that this value is not assumed to represent

a similar air-entry pressure for all soils; rather, it serves as a conservative lower bound defining fully saturated conditions across all soils.

For implementing this constraint, as illustrated in Fig. 2b, we generate a set of residual points, which are randomly sampled within the input space using *pF* values between −2.0 and −0.3. During training, we evaluate Eq. (10) at these points and penalize deviations from this constraint by adding the fourth term to the loss function (Eq. (9)). The numbers of residual points used to enforce these two physical constraints, denoted as S_1 and S_2 in Eq. (9), were selected through manual tuning.

2.1.7. Model training via automatic differentiation

The hybrid model developed in this study involves several interconnected neural networks, each containing trainable parameters (Fig. 1). These networks are coupled in Eq. (8) through a physics-constrained ansatz (i.e., a prior mathematical form assumed to guide the model). This coupling of various neural networks results in a highly nonlinear system whose parameters must be optimized to minimize the total loss function defined in Eq. (9).

Training such a hybrid system requires computing gradients of the loss function, Eq. (9), with respect to all trainable parameters. To enable efficient and accurate gradient computation, we leverage automatic differentiation (AD), a core feature in modern deep learning frameworks such as TensorFlow and PyTorch (Baydin et al., 2018). Automatic differentiation automatically constructs a computational graph during the model's forward pass and traces the sequence of mathematical operations from inputs to outputs. During backpropagation, reverse-mode AD traverses this graph from the output layer back to the inputs and systematically applies the chain rule to compute exact gradients with respect to every trainable parameter. This allows the model to be trained efficiently using standard gradient-based optimization algorithms, despite its architectural complexity and the presence of embedded physical constraints.

2.1.8. Hyperparameter optimization

The developed hybrid model consists of separate sub-network neural models, each designed for a specific component. Table 1 reports the inputs, architecture, and activation functions of each sub-network. To select the optimal architecture of the interior networks, we initially adopted relatively expressive network structures and subsequently reduced the number of layers and units until a noticeable degradation in model performance was observed. The final architectures were chosen as the simplest configurations that maintained stable performance. All hidden layers across the sub-networks employed Exponential Linear Unit (ELU) activation functions.

The output of networks NN_c, NN_s, NN_{dry}, and NN_o were implemented with sigmoid activation functions and were scaled appropriately to their respective physical ranges. Additionally, the networks NN_s, NN_o, and NN_{dry}, which estimate soil-dependent constants, shared a similar structure: two hidden layers with four units each and ELU activations.

We set λ_1 and λ_2 in Eq. (9) to 1 and 12.1, respectively, based on manual tuning. The parameters λ_3 and λ_4 were set to 10 and 5, respectively. These weights were tuned to ensure their corresponding constraints were satisfied without degrading the overall performance of the model. The parameters S_1 and S_2, which determine the number of residual points in sets 1 and 2, were both set to 2000.

The Adam optimizer (Kingma, 2014) was used for model training with an initial learning rate of 0.005. An adaptive learning rate strategy was applied, in which the learning rate was reduced by a factor of 0.8 if no improvement in the validation loss was observed, continuing down to a minimum of 0.0005.

To avoid overfitting, early stopping with a patience of 10 epochs was applied. In practice, this resulted in model convergence after approximately 250 training epochs. The evolution of the training and validation loss as a function of epoch is provided in the Supplementary Information (Fig. S1). Additionally, because the capillary sub-network (NN_c) directly influences the smoothness of the final model, we applied L2 regularization with an intensity of 0.15 to all layers of this sub-network. L2 regularization prevents overfitting by controlling the magnitude of large weights in the network (Ng, 2004). The model was entirely developed in Python and implemented using TensorFlow (Abadi et al., 2016).

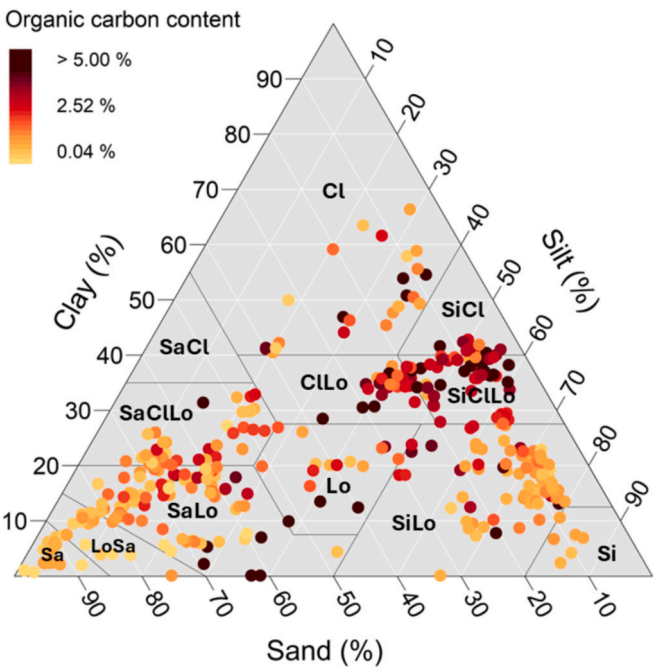

Fig. 3. The textural distribution of 482 soil samples according to the USDA (United States Department of Agriculture) soil triangle. Point colors indicate organic carbon content (%) of each soil sample. The abbreviations Sa, Si, and Cl stand for sand, silt, and clay, respectively.

Table 2
Summary statistics of soil physical properties in the dataset.

	Sand [%]	Silt [%]	Clay [%]	OC [%]	BD [gr cm^{-3}]
Min	3.3	0.0	0.1	0.04	0.37
Mean	41.3	38.6	20.1	2.11	1.33
Max	99.9	85.7	66.4	19.26	1.89
Std	30.4	24.6	13.2	2.15	0.29

OC, organic carbon; BD, bulk density; Std, standard deviation

2.2. Experimental data and train/validation/test splits

For training and evaluating the hybrid pedotransfer functions in this study, we used 482 undisturbed soil samples from the publicly available dataset of Hohenbrink et al. (2023), which includes measurements of soil hydraulic properties for a wide range of texture types and organic carbon contents (Fig. 3 and Table 2). The 482 soils were extracted from the main dataset (consisting of 572 soil samples) with the condition that measurements of SWRC, soil textural fractions, bulk density, and organic carbon content, are available. This dataset is available at doi:10.5880/fidgeo.2023.012.

As with all data-driven models, predictive performance is strongly dependent on the dataset size (Ahmadisharaf et al., 2024) and diversity of input training examples. The present dataset was selected to maximize both sample size as well as diversity in soil texture, organic carbon content, and coverage of both wet- and dry-end measurements, which is essential for learning representative intermediate retention processes.

Particle size distribution was measured using wet sieving and sedimentation techniques. The size classes were categorized following the

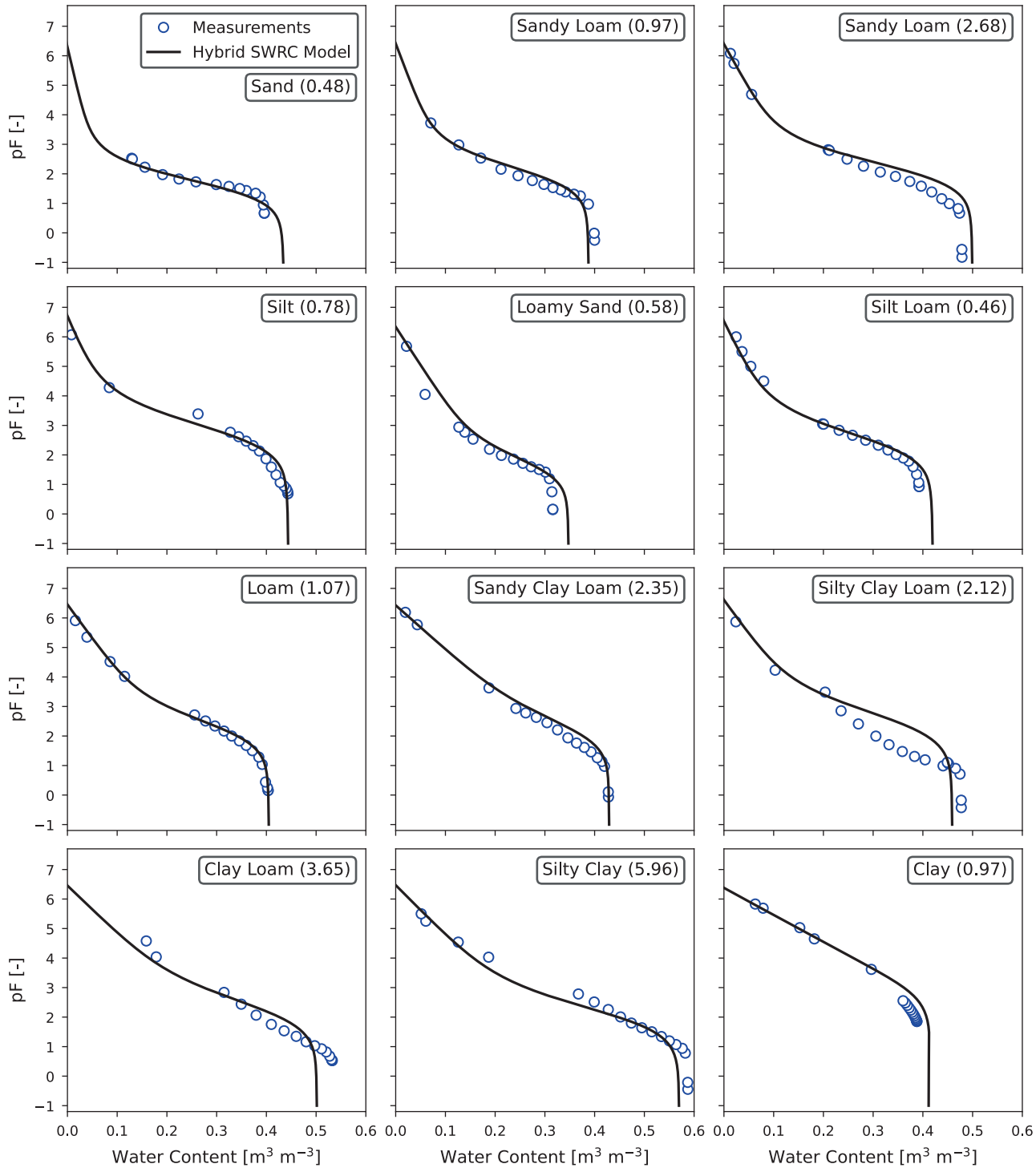

Fig. 4. Predicted soil water retention curves from the developed hybrid pedotransfer functions for twelve soils representing different USDA textural classes. Values in parentheses represent the organic carbon content (OC) in percentage.

United States Department of Agriculture (USDA) classification system, which defines clay as particles smaller than 2 μm, silt as 2–50 μm, and sand as 50–2,000 μm (Hillel, 1982). The selected set of soil samples covers eleven soil textural classes of USDA system, making the dataset highly suitable for the data-driven approach in this study (Fig. 3).

Organic carbon content was measured by high-temperature combustion using an elemental analyzer. Bulk density was determined gravimetrically, by oven-drying the soil samples for at least 24 h following evaporation experiments (Tehrani et al., 2023).

The dataset includes measurements of the soil water retention curve covering a broad range of matric heads. The wet and medium moisture ranges ($pF \leq 4.2$) were measured using the simplified evaporation method (Peters and Durner, 2008; Schindler, 1980), implemented via the HYPROP device (METER Group AG, Germany). This method captures the drying branch of the retention curve and provides high-resolution measurements. For the dry end ($pF > 4.2$), additional

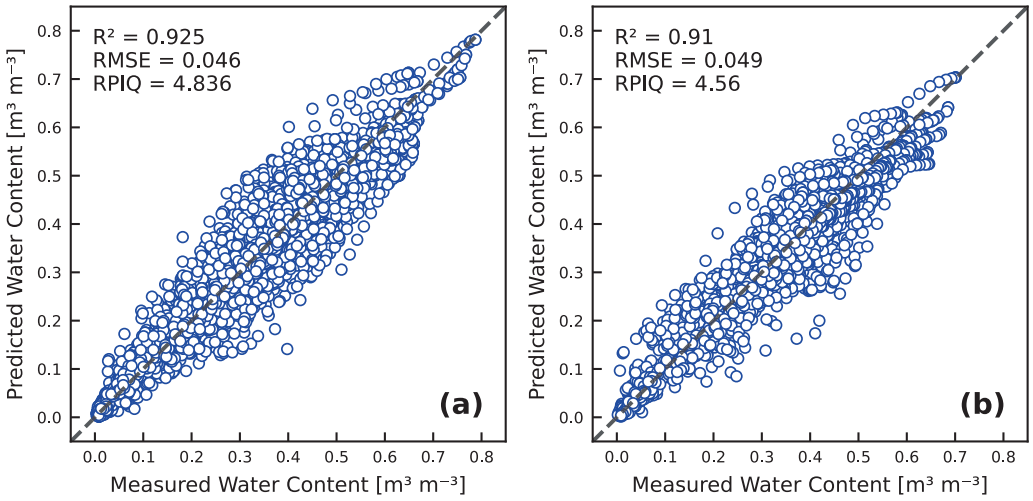

Fig. 5. Overall performance of the hybrid pedotransfer functions on the (a) train and (b) test sets. Validation performance metrics are RMSE = 0.045 $m^3\ m^{-3}$, R^2 = 0.924, and RPIQ = 4.857. The validation set, comprising 20 % of the training data, was used during training for model selection and early stopping.

measurements were obtained using the dewpoint method (Campbell et al., 2007; Kirste et al., 2019) via the WP4C device (METER Group, Inc., USA). Due to the large imbalance between the number of dry-end and wet-end data points, we uniformly resampled the wet-end measurements to 15 points. For more detailed information about the measurement details and device specifications, readers are referred to Hohenbrink et al. (2023).

To train the model with the customized loss function in Eq. (9), which includes separate terms for the wet and dry ends of the SWRC, we partitioned the data for each end into training and test sets separately to achieve a more balanced split. Seventy percent of the samples from each end were used for training and validation, while the remaining 30 % served as a hold-out test set. The stratified sampling ensured that each subset, training, validation, and testing, represented the full range of wet and dry conditions. Within the 70 % training pool, 20 % of the data were further set aside as a validation set during training. The dataset comprised 8394 $pF-\theta$ observations after resampling, of which 5916 $pF-\theta$ pairs (339 samples) were used for training and validation, and 2478 $pF-\theta$ pairs (143 samples) for testing.

To avoid data leakage and account for the strong correlation among water content measurements within individual soil samples, we partitioned the dataset at the soil sample level rather than at the level of individual $pF-\theta$ pairs. This ensured that all $pF-\theta$ observations from a given sample were assigned entirely to either the training or testing set. Additionally, we ensured that samples from the same soil profile (i.e., location) were exclusively included in either the training or the testing set.

2.3. Evaluation criteria

Model performance was evaluated based on three metrics: root mean square error (RMSE), coefficient of determination (R-squared), and the ratio of the interquartile range to RMSE (RPIQ), all calculated using the predicted and observed volumetric water contents:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\theta}^{(i)} - \theta^{(i)}\right)^2} \tag{11}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(\widehat{\theta}^{(i)} - \theta^{(i)}\right)^2}{\sum_{i=1}^{N}\left(\theta^{(i)} - \overline{\theta}\right)^2} \tag{12}$$

$$RPIQ = \frac{Q_{75} - Q_{25}}{RMSE} \tag{13}$$

where N is the total number of measured points, $\overline{\theta}$ is the mean of the measured water contents, and Q_{75} and Q_{25} correspond to the 75th and 25th percentiles of the measured water contents, respectively. The RPIQ metric offers a scale-independent metric by comparing the RMSE with the variability of the data.

3. Results and discussion

3.1. Physics-constrained neural network performance

Fig. 4 shows the predicted shape of the SWRC for twelve soil samples with different texture classes according to the USDA classification system. The model is trained by optimizing the loss function (Eq. (9)) over the entire training set as a whole, rather than fitting it sample by sample. Unlike parametric PTFs, which use predefined analytical forms for the SWRC, our hybrid model learns the curve shape directly from the data. Once trained, the model can predict the entire continuous SWRC. To achieve this, for any given soil with fixed physical properties (i.e., sand, silt, clay, OC, and BD), we vary the pF over a specified range to generate the continuous curves shown in Fig. 4.

The discovered shapes of the SWRCs are smooth, differentiable, and therefore suitable for simulation of soil water flow (i.e., Richardson-Richards equation). These curves exhibit a sigmoidal shape in the wet range, similar to traditional parametric models (van Genuchten, 1980; Kosugi, 1994), and transition to a linear form at lower water contents, consistent with the Campbell-Shiozawa model behavior assumed at the dry end. Notably, the transition between the neural network-predicted region and the analytically modeled region governed by the Campbell-Shiozawa model is seamless, with no noticeable discontinuities or abrupt changes, resulting in a smooth, continuous curve. Furthermore, at the wet end, the curves remain invariant with respect to pF for values below −0.3. At the dry end, the range of pF at zero water content, pF_{dry}, for all curves remains between 6.2 and 7.6, ensuring the satisfaction of the physical constraints imposed at both ends during training.

The overall performance of the hybrid PTFs on both the train and test sets is depicted in Fig. 5. As shown, the model demonstrates a close performance on both the train and test sets, indicating the generalization capability of the model. This close performance between training and testing sets is particularly important in developing continuous, non-

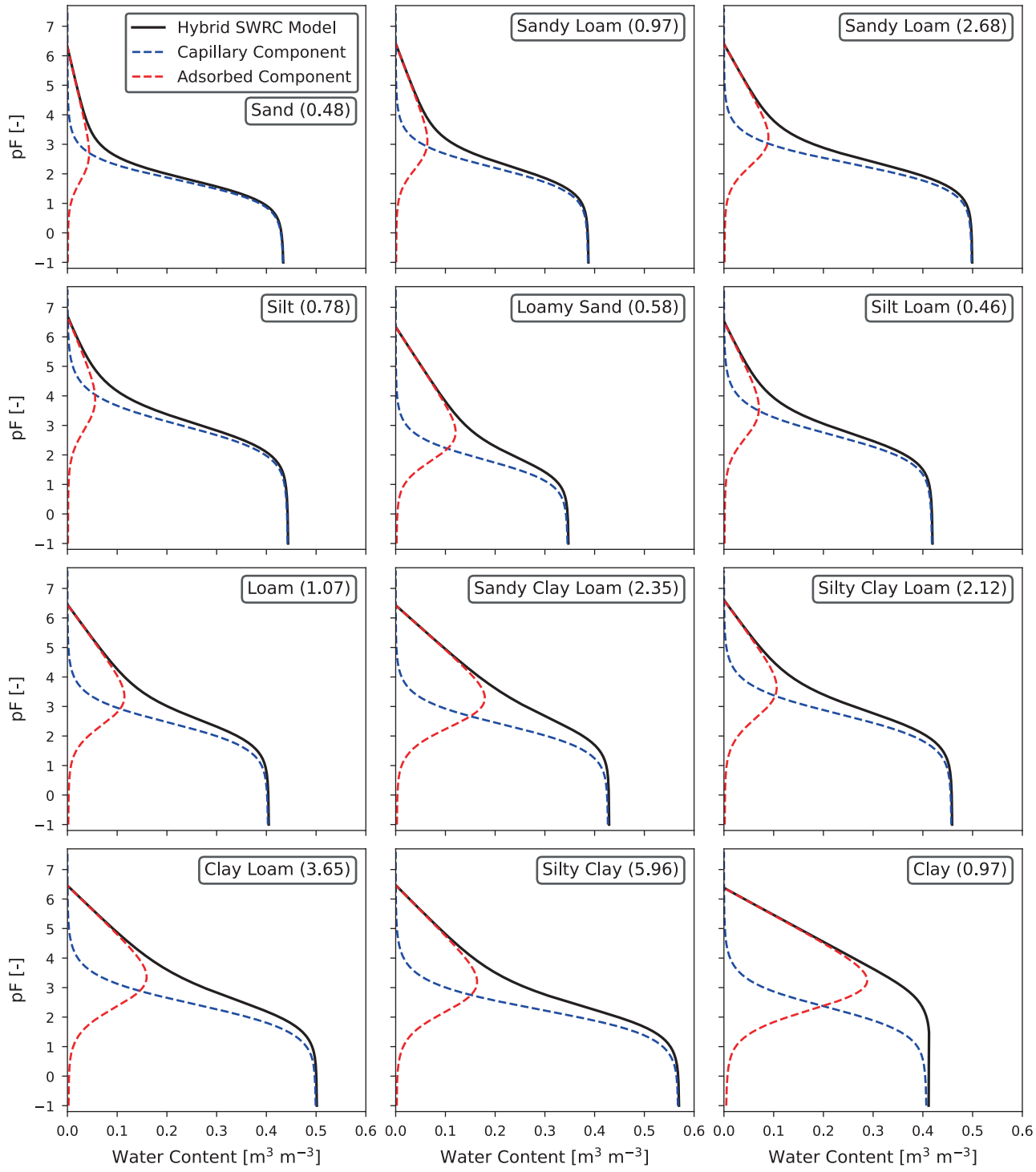

Fig. 6. Discovered capillary and adsorbed film components of the soil water retention curve for the same soil samples shown in Fig. 4, as obtained from the hybrid model. Values in parentheses represent the organic carbon content (OC) in percentage.

parametric and semi-parametric PTFs, as even small degrees of overfitting can distort the predicted curve, making it physically unrealistic.

The model achieved an RMSE of 0.049 $m^3\ m^{-3}$ on the test set, which is reasonable given the diversity of soils represented in the dataset, including eleven USDA texture classes, and samples with high organic carbon content and very low bulk density (Table 2). The extent of variations in soil properties is also reflected in Fig. 5, where measured volumetric water contents reach values as high as 0.8 $m^3\ m^{-3}$. This diversity, along with dataset size, variations in soil properties, and measurement quality, are key factors influencing the performance of PTFs. The performance of the PTF developed aligns well with both continuous parametric models (dos Santos Pereira et al., 2025) and continuous non-parametric models trained on HYPROP system measurements (Haghverdi et al., 2018). However, due to differences in the datasets used and input predictors across studies, direct comparisons are challenging.

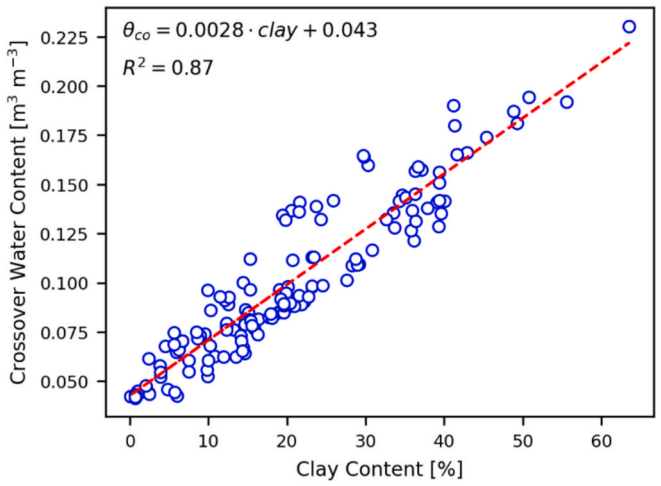

Fig. 7. Water content at the crossover point between capillary and adsorbed film components, denoted as θ_{co}, versus clay content for all soils in the test set. In the regression equation shown in the figure, the variable 'clay' represents the clay content (%).

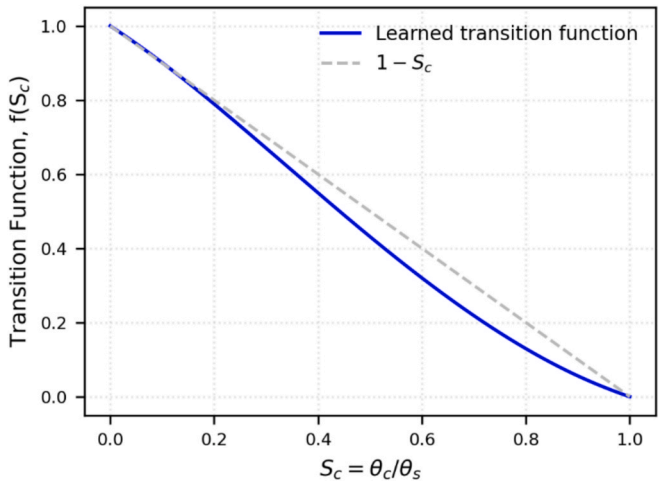

Fig. 8. Learned transition function, $f(S_c)$, as defined in Eq. (5). The reference curve $1 - S_c$, commonly assumed in previous studies, is shown in gray to illustrate the deviation of the learned function from linear behavior.

The discovered partitioning of capillary and adsorbed film components of the SWRC for the soil samples shown in Fig. 4 is depicted in Fig. 6. These curves are obtained by plotting the first and second terms on the right-hand side of Eq. (8). The curves shown in Fig. 6 should be considered as follows: at each *pF*, the total water content (black line) is the sum of the capillary and adsorbed film water contents.

This data-driven partitioning aligns remarkably well with the physics-based models in the literature, such as that of Or and Tuller (1999) model, which was developed by incorporating detailed interfacial physics within an angular pore geometry. Specifically, the capillary component dominates under saturated conditions for all soils. This corresponds to the point where the liquid–vapor interfacial area is effectively zero. As *pF* increases (corresponding to more suction in soil), pores of varying sizes begin to drain, and this process starts with larger pores. For each pore size, there exists a critical *pF*, often referred to as the "onset of drainage", at which air starts to invade the pore.

As drainage progresses, water films begin to form along the surfaces of the partially emptied pores. With further increase of *pF*, smaller pores also undergo drainage, leading to a gradual decrease in the capillary component and a concurrent increase in the contribution of the film component. This reflects the physical process by which surface area becomes increasingly available for film water as the capillary water recedes to the pores corners (Or and Tuller, 1999).

At a *pF* between 2.5 and 4 for all soils, the film component reaches a peak. Beyond this peak, toward the dry region, both capillary and adsorbed film contributions decline; this is the point where even the adsorbed water films begin to thin. This decline continues at very high *pF* values and eventually the adsorbed water converges to the Campbell-Shiozawa linear model for the dry end, with the extent of this linear region being dependent on soil texture (Fig. 6). Note that, consistent with physical constraints imposed, the water content for $pF > 5$ remains entirely in adsorbed form for all soils.

Another important point in Fig. 6 is the crossover between capillary and adsorbed component curves, which determines the boundary between capillary and adsorbed dominant regions. This crossover point is highly dependent on soil texture, and as seen in Figs. 6 and 7, with an increase in clay content, the water content at this point increases. This is because the fraction of finer particles (i.e., silt and clay) provides a greater specific surface area, which supports the formation and retention of more extensive water films along particle surfaces (Norouzi et al., 2022; Norouzi et al., 2023).

An important advantage of DHM is the ability to replace unknown or poorly defined components with neural networks, which serve as universal function approximators. In our model, the transition function $f(S_c)$ in Eq. (5), whose analytical form was not known in advance, was learned directly from data. This function governs how the Campbell-Shiozawa model should be modified at *pF* values lower than a certain threshold, where capillary and adsorbed film water may coexist. This learned function is shown in Fig. 8. As illustrated, $f(S_c)$ exhibits a decreasing trend, which is expected: as capillary water increases, the contribution of the Campbell-Shiozawa model to total water content should diminish.

The transition function was designed with a hard constraint to satisfy $f(S_c = 0) = 1$, ensuring full reliance on the Campbell-Shiozawa model when the capillary saturation is zero (Fig. 8). Interestingly, although not explicitly constrained to do so, the learned function also satisfies $f(S_c = 1) = 0$. This behavior implies that at capillary saturation ($S_c = 1$), the entire water content is attributed to capillary water, with no contribution from the adsorbed film component (Fig. 6). Additionally, the transition between the two endpoints ($S_c = 0$ and $S_c = 1$) is distinctly nonlinear, deviating from the commonly assumed linear transition ($1 - S_c$) used in previous studies (Fayer and Simmons, 1995; Lebeau and Konrad, 2010).

3.2. Performance of the sub-networks for soil constants

One of the key advantages of DHM is its ability to train multiple internal models simultaneously within a unified framework (Shen et al., 2023). In our hybrid PTF, we implemented three dedicated sub-networks (i.e., sub-PTFs) to estimate key soil parameters, namely, θ_s, pF_{dry}, and θ_o, directly from basic soil physical properties (see Eq. (8) and Table 1). Each of these sub-networks maps soil properties to a physical constant. After training, each of these networks could be used as a standalone PTF. It should be noted that these sub-networks were not trained with separate target data; instead, they adjusted their parameters as part of the joint training of the full hybrid model (Fig. 1).

Fig. 9 illustrates the predictions of each sub-network, plotted against one of the representative input variables, to ensure that these sub-networks have learned meaningful physically or empirically known relationships rather than overfitting the training data.

The sub-network NN_s predicts θ_s. The predictions of this sub-network exhibit a clear inverse relationship with bulk density in its input layer, consistent with the physical understanding that higher bulk density typically corresponds to lower total porosity and thus lower saturated water content (Fig. 9a). Interestingly, for high bulk density values, the predictions align well with the standard equation used to calculate

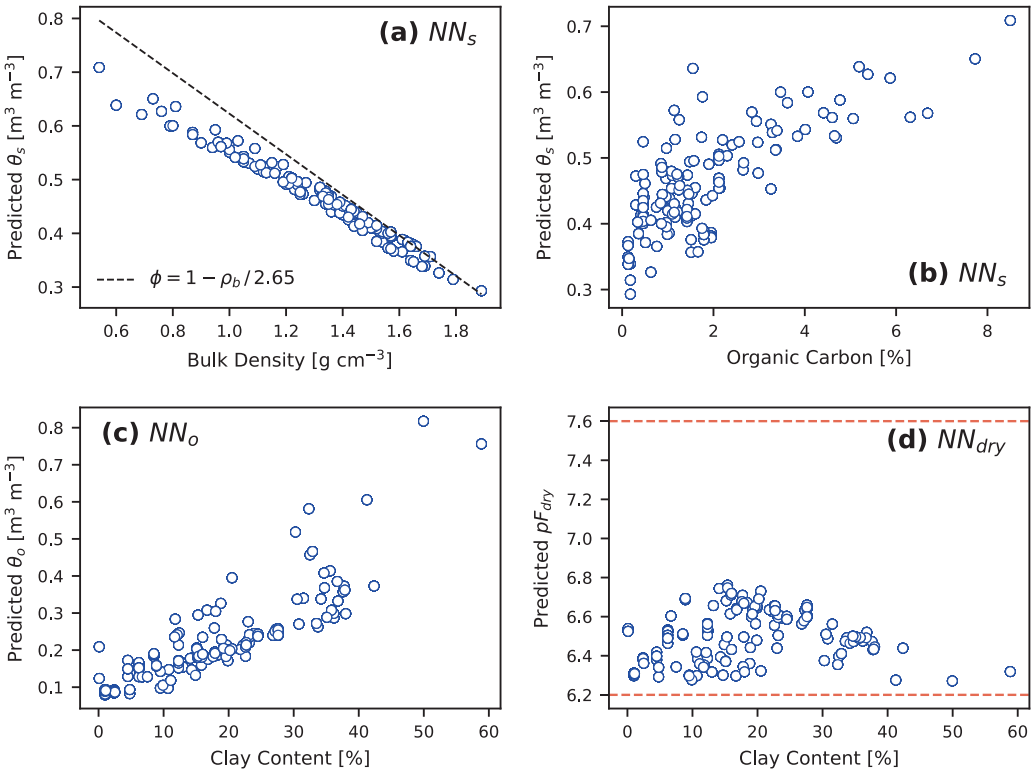

Fig. 9. Predicted soil-specific parameters from the interior sub-networks of the hybrid PTF model, plotted against representative input variables. The parameters θ_s and ϕ represent saturated water content and porosity, respectively, and θ_o and pF_{dry} are the fitting parameters that determine the slope and intercept of the Campbell-Shiozawa model at the dry end [Eq. (2)].

porosity from known bulk density, assuming a particle density of 2.65 g/cm^3. However, for soils with low bulk density (organic soils), the predictions deviate from this relationship, demonstrating the inadequacy of the 2.65 g/cm^3 assumption for such soils (Marakkala Manage et al., 2019).

In Fig. (9b) the predictions of NN_s are plotted against OC in the input layer. As expected, saturated water content increases with increasing OC. Organic particles have lower intrinsic density with more irregular, often fibrous structures than mineral particles, which leads to looser packing and greater total pore volume. Consequently, soils with higher organic carbon content tend to exhibit higher porosity and lower bulk density, consistent with observations in analytical modeling studies (compare for example with the trend observed in Fig. 4 of Robinson (2022)).

Similarly, the sub-network NN_o (Fig. 9c), which predicts θ_o of Campbell-Shiozawa model and determines the slope of the linear region at the dry end, shows a positive correlation with clay content. As noted by Campbell and Shiozawa (1992), this free parameter is highly influenced by the amount of soil clay content, which determines the specific area of soil. Higher θ_o values are indicative of finer-textured soils, which retain more water across a wide range of matric heads.

In contrast to the other two sub-networks, the sub-network predicting pF_{dry}, denoted as NN_{dry}, does not show any clear correlation with the input variables. This is consistent with findings by Lu and Khorshidi (2015) and Karup et al. (2017), who showed that pF_{dry} is more dependent on clay mineralogy than soil OC or clay content. Although no strong correlation is observed, all predictions from this sub-network remain within the imposed limits ($pF = 6.2$ to 7.6), which were enforced as a hard constraint by scaling the network's output to the target range (Fig. 9d).

These results collectively demonstrate that the interior sub-networks not only remained within the specified physical bounds but also captured relationships that are consistent with established physical and empirical understanding.

4. Conclusions and outlook

We draw the following main conclusions from this study:

1) We presented a differentiable hybrid modeling (DHM) framework that combines mechanistic understanding with data-driven components to discover internal and largely unobservable soil processes that are not directly measurable. Applied to the soil water retention curve, the hybrid approach successfully learned the overall shape of the SWRC as well as its capillary and adsorbed components. Notably, the hybrid model learned pore-scale features without relying on explicit geometrical assumptions about soil pore shape or its distribution.
2) Our model demonstrates a new perspective on the use of data in soil physics. We used the same inputs and outputs as conventional pedotransfer functions but our main objective goes far beyond simple prediction. During this mapping from inputs to outputs, the hybrid model learns multiple intermediate processes and their unobservable transitions implicitly, without requiring explicit data for them. Importantly, these learned internal relationships produced physically meaningful results and as observed in the case of transition function, the model discovered a nonlinear function that challenges the linear assumptions invoked in previous studies.
3) The DHM framework is flexible in design. As our understanding of soil physical processes advances, we can incorporate more soil physics knowledge into the hybrid model structure. Future work may explore alternative formulations for dry-end behavior and evaluate how learned sub-models can be reused in broader hydrological modeling contexts.

Given the increasing availability of large soil datasets, we believe DHM and its capability for end-to-end training of several internal components (i.e., sub-models) within a single optimization process, provides a promising tool that can be leveraged for modeling fundamental physical processes where partial knowledge of the underlying mechanisms has led to over-simplifying assumptions and biased predictions.

Declaration of Generative AI in the writing process

In the preparation of this paper, ChatGPT 5 was used to assist with improving the clarity, grammar, and overall quality of the English language. Following its use, the authors thoroughly reviewed and edited the content to ensure accuracy and alignment with the intended meaning. The AI's involvement was limited to language refinement, and no content generation, interpretation of data, or intellectual contributions were made by the AI. The responsibility for the scientific content, and conclusions of this paper lies entirely with the authors.

CRediT authorship contribution statement

Sarem Norouzi: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Formal analysis, Conceptualization. **Per Moldrup:** Writing – review & editing, Methodology, Conceptualization. **Ben Moseley:** Writing – review & editing, Methodology, Conceptualization. **David Robinson:** Writing – review & editing, Methodology, Conceptualization. **Dani Or:** Writing – review & editing, Methodology, Conceptualization. **Tobias L. Hohenbrink:** Writing – review & editing, Visualization, Formal analysis, Data curation. **Budiman Minasny:** Writing – review & editing, Methodology. **Morteza Sadeghi:** Writing – review & editing, Methodology. **Emmanuel Arthur:** Writing – review & editing, Methodology. **Markus Tuller:** Writing – review & editing, Conceptualization. **Mogens H. Greve:** Writing – review & editing, Supervision, Resources, Methodology, Funding acquisition. **Lis W. de Jonge:** Writing – review & editing, Supervision, Resources, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors gratefully acknowledge the late Professor Markus Tuller, whose mentorship and insightful collaborations greatly influenced the ideas and results presented in this work.

This study is funded by the European Union Horizon Europe research and innovation program under grant agreement No. 101086179 (AI4SoilHealth). Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2026.135008.

Data availability

The soil water retention data used in this study are publicly available and were obtained from the dataset published by Hohenbrink et al. (2023).

The code used to implement the differentiable hybrid modeling framework and to reproduce the results presented in this study will be made available by the authors upon request.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mane, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viegas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X., 2016. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. doi:10.48550/arXiv.1603.04467.

Ahmadisharaf, A., Nematirad, R., Sabouri, S., Pachepsky, Y., Ghanbarian, B., 2024. Representative sample size for estimating saturated hydraulic conductivity via machine learning: a proof-of-concept study. Water Resour. Res. 60, e2023WR036783. https://doi.org/10.1029/2023WR036783.

Arthur, E., Tuller, M., Moldrup, P., Resurreccion, A.C., Meding, M.S., Kawamoto, K., Komatsu, T., de Jonge, L.W., 2013. Soil specific surface area and non-singularity of soil-water retention at low saturations. Soil Sci. Soc. Am. J. 77, 43–53. https://doi.org/10.2136/sssaj2012.0262.

Bandai, T., Ghezzehei, T.A., 2021. Physics-informed neural networks with monotonicity constraints for Richardson-Richards equation: estimation of constitutive relationships and soil water flux density from volumetric water content measurements. Water Resour. Res. 57, e2020WR027642. https://doi.org/10.1029/2020WR027642.

Baydin, A.G., Pearlmutter, B.A., Radul, A.A., Siskind, J.M., 2018. Automatic differentiation in machine learning: a survey. J. Machine Learning Res. https://doi.org/10.48550/arXiv.1502.05767.

Brusseau, M.L., 2023. Determining air-water interfacial areas for the retention and transport of PFAS and other interfacially active solutes in unsaturated porous media. Sci. Total Environ. 884, 163730. https://doi.org/10.1016/j.scitotenv.2023.163730.

Campbell, G.S., Shiozawa, S., 1992. Prediction of hydraulic properties of soils using particle-size distribution and bulk density data. In: International Workshop on Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils. University of California, California, pp. 317–328.

Campbell, G.S., Smith, D.M., Teare, B.L., 2007. Application of a dew point method to obtain the soil water characteristic. In: Schanz, T. (Ed.), Experimental Unsaturated Soil Mechanics. Springer, Berlin, Heidelberg, pp. 71–77. https://doi.org/10.1007/3-540-69873-6_7.

Derjaguin, B.V., Churaev, N.V., Muller, V.M., 1987. Surface Forces. Springer US, Boston, MA. doi:10.1007/978-1-4757-6639-4.

dos Santos Pereira, S.A., de F.N. Gitirana, G., Mendes, T.A., de A. Gomes, R., 2025. Artificial neural networks for the prediction of the soil-water characteristic curve: an overview. Soil Tillage Res. 248, 106466. https://doi.org/10.1016/j.still.2025.106466.

Fayer, M.J., Simmons, C.S., 1995. Modified soil water retention functions for all matric suctions. Water Resour. Res. 31, 1233–1238. https://doi.org/10.1029/95WR00173.

Fredlund, D.G., Xing, A., 1994. Equations for the soil-water characteristic curve. Can. Geotech. J. 31, 521–532. https://doi.org/10.1139/t94-061.

Gardner, W., 1958. Some steady-state solutions of the unsaturated moisture flow equation with application to evaporation from a water table. Soil Sci. 85, 228–232.

Ghorbani, A., Babaeian, E., Sadeghi, M., Durner, W., Jones, S.B., van Genuchten, M.T., 2025. An improved van Genuchten soil water characteristic model to account for surface adsorptive forces. J. Hydrol., 133692 https://doi.org/10.1016/j.jhydrol.2025.133692.

Green, W.H., Ampt, G.A., 1911. Studies on soil phyics. J. Agric. Sci. 4, 1–24. https://doi.org/10.1017/S0021859600001441.

Guo, B., Zeng, J., Brusseau, M.L., 2020. A mathematical model for the release, transport, and retention of Per- and Polyfluoroalkyl Substances (PFAS) in the Vadose Zone. Water Resour. Res. 56, e2019WR026667. https://doi.org/10.1029/2019WR026667.

Gupta, S., Papritz, A., Lehmann, P., Hengl, T., Bonetti, S., Or, D., 2022. Global mapping of soil water characteristics parameters— fusing curated data with machine learning and environmental covariates. Remote Sens. (Basel) 14, 1947. https://doi.org/10.3390/rs14081947.

Haghverdi, A., Cornelis, W.M., Ghahraman, B., 2012. A pseudo-continuous neural network approach for developing water retention pedotransfer functions with limited data. J. Hydrol. 442–443, 46–54. https://doi.org/10.1016/j.jhydrol.2012.03.036.

Haghverdi, A., Öztürk, H.S., Durner, W., 2018. Measurement and estimation of the soil water retention curve using the evaporation method and the pseudo continuous pedotransfer function. J. Hydrol. 563, 251–259. https://doi.org/10.1016/j.jhydrol.2018.06.007.

Haruzi, P., Moreno, Z., 2023. Modeling water flow and solute transport in unsaturated soils using physics-informed neural networks trained with geoelectrical data. Water Resour. Res. 59, e2023WR034538. https://doi.org/10.1029/2023WR034538.

Hillel, D., 1982. 3 - Texture, particle size distribution, and specific surface. In: Hillel, D. (Ed.), Introduction to Soil Physics. Academic Press, San Diego, pp. 21–39. https://doi.org/10.1016/B978-0-08-091869-3.50007-X.

Hohenbrink, T., Jackisch, C., Durner, W., Germer, K., Iden, S., Kreiselmeier, J., Leuther, F., Metzger, J., Naseri, M., Peters, A., 2023. Soil Water retention and hydraulic conductivity measured in a wide saturation range. Earth Syst. Sci. Data. https://doi.org/10.5194/essd-2023-74.

Karniadakis, G., Kevrekidis, Y., Lu, L., Perdikaris, P., Wang, S., Yang, L., 2021. Physics-informed machine learning. Nat. Rev. Phys. 1–19. https://doi.org/10.1038/s42254-021-00314-5.

Karup, D., Moldrup, P., Tuller, M., Arthur, E., De Jonge, L.W., 2017. Prediction of the soil water retention curve for structured soil from saturation to oven-dryness. European J Soil Science 68, 57–65. https://doi.org/10.1111/ejss.12401.

Kingma, D.P., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv: 1412.6980.
Kirste, B., Iden, S.C., Durner, W., 2019. Determination of the soil water retention curve around the wilting point: optimized protocol for the dewpoint method. Soil Sci. Soc. Am. J. 83, 288–299. https://doi.org/10.2136/sssaj2018.08.0286.
Kosugi, K., 1994. Three-parameter lognormal distribution model for soil water retention. Water Resour. Res. 30, 891–901. https://doi.org/10.1029/93WR02931.
Lebeau, M., Konrad, J.-M., 2010. A new capillary and thin film flow model for predicting the hydraulic conductivity of unsaturated porous media. Water Resour. Res. 46. https://doi.org/10.1029/2010WR009092.
Lu, N., 2016. Generalized soil water retention equation for adsorption and capillarity. J. Geotech. Geoenviron. Eng. 142, 04016051. https://doi.org/10.1061/(ASCE)GT.1943-5606.0001524.
Lu, N., Khorshidi, M., 2015. Mechanisms for soil-water retention and hysteresis at high suction range. J. Geotech. Geoenviron. Eng. 141, 04015032. https://doi.org/10.1061/(ASCE)GT.1943-5606.0001325.
Marakkala Manage, L.P., Katuwal, S., Norgaard, T., Knadel, M., Moldrup, P., De Jonge, L. W., Greve, M.H., 2019. Estimating Soil particle density using visible near-infrared spectroscopy and a simple, two-compartment pedotransfer function. Soil Science Soc of Amer J 83, 37–47. https://doi.org/10.2136/sssaj2018.06.0217.
Minasny, B., Bandai, T., Ghezzehei, T.A., Huang, Y.-C., Ma, Y., McBratney, A.B., Ng, W., Norouzi, S., Padarian, J., Rudiyanto Sharififar, A., Styc, Q., Widyastuti, M., 2024. Soil Science-Informed Machine Learning. Geoderma 452, 117094. https://doi.org/10.1016/j.geoderma.2024.117094.
Moseley, B., 2022. Physics-informed machine learning: from concepts to real-world applications. University of Oxford (PhD Thesis).
Nachum, S., 2025. Soil water potential in geosciences: an overview. Geosciences 15, 123. https://doi.org/10.3390/geosciences15040123.
Ng, A.Y., 2004. Feature selection, L 1 vs. L 2 regularization, and rotational invariance. In: Presented at the Proceedings of the Twenty-First International Conference on Machine Learning, p. 78.
Norouzi, S., Pesch, C., Arthur, E., Norgaard, T., Moldrup, P., Greve, M.H., Beucher, A.M., Sadeghi, M., Zaresourmanabad, M., Tuller, M., Iversen, B.V., de Jonge, L.W., 2025. Physics-informed neural networks for estimating a continuous form of the soil water retention curve from basic soil properties. Water Resour. Res. 61, e2024WR038149. https://doi.org/10.1029/2024WR038149.
Norouzi, S., Sadeghi, M., Tuller, M., Ebrahimian, H., Liaghat, A., Jones, S.B., De Jonge, L. W., 2023. A novel laboratory method for the retrieval of the soil water retention curve from shortwave infrared reflectance. J. Hydrol. 626, 130284. https://doi.org/10.1016/j.jhydrol.2023.130284.
Norouzi, S., Sadeghi, M., Tuller, M., Liaghat, A., Jones, S.B., Ebrahimian, H., 2022. A novel physical-empirical model linking shortwave infrared reflectance and soil water retention. J. Hydrol. 614, 128653. https://doi.org/10.1016/j.jhydrol.2022.128653.
Or, D., Tuller, M., 1999. Liquid retention and interfacial area in variably saturated porous media: upscaling from single-pore to sample-scale model. Water Resour. Res. 35, 3591–3605. https://doi.org/10.1029/1999WR900262.
Peters, A., 2013. Simple consistent models for water retention and hydraulic conductivity in the complete moisture range: hydraulic models for the complete moisture range. Water Resour. Res. 49, 6765–6780. https://doi.org/10.1002/wrcr.20548.
Peters, A., Durner, W., 2008. Simplified evaporation method for determining soil hydraulic properties. J. Hydrol. 356, 147–162. https://doi.org/10.1016/j.jhydrol.2008.04.016.
Philip, J.R., de De Vries, D., 1957. Moisture movement in porous materials under temperature gradients. Eos Trans. AGU 38, 222–232.
Psichogios, D.C., Ungar, L.H., 1992. A hybrid neural network-first principles approach to process modeling. AIChE J 38, 1499–1511. https://doi.org/10.1002/aic.690381003.
Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J. Comput. Phys. 378, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045.
Rasoulzadeh, A., Bezaatpour, J., Azizi Mobaser, J., Fernández-Gálvez, J., 2025. Half-century review and advances in closed-form functions for estimating soil water retention curves. Hydrology 12, 164. https://doi.org/10.3390/hydrology12070164.
Rawls, W.J., Brakensiek, D.L., Saxtonn, K.E., 1982. Estimation of soil water properties. Transactions of the ASAE 25, 1316–1320. https://doi.org/10.13031/2013.33720.
Robinson, D.A., 2022. Analytical modelling of soil porosity and bulk density across the soil organic matter and land-use continuum. Sci. Rep.
Schindler, U., 1980. Ein schnellverfahren zur messung der wasserleitfähigkeit im teilgesättigten boden an stechzylinderproben.
Schneider, M., Goss, K.-U., 2012. Prediction of the water sorption isotherm in air dry soils. Geoderma 170, 64–69. https://doi.org/10.1016/j.geoderma.2011.10.008.
Shen, C., Appling, A.P., Gentine, P., Bandai, T., Gupta, H., Tartakovsky, A., Baity-Jesi, M., Fenicia, F., Kifer, D., Li, L., Liu, X., Ren, W., Zheng, Y., Harman, C.J., Clark, M., Farthing, M., Feng, D., Kumar, P., Aboelyazeed, D., Rahmani, F., Song, Y., Beck, H.E., Bindas, T., Dwivedi, D., Fang, K., Höge, M., Rackauckas, C., Mohanty, B., Roy, T., Xu, C., Lawson, K., 2023. Differentiable modelling to unify machine learning and physical models for geosciences. Nat. Rev. Earth Environ. 4, 552–567. https://doi.org/10.1038/s43017-023-00450-9.
Sourmanabad, M.Z., Norouzi, S., Mirzaei, F., Yokeley, B.A., Ebrahimian, H., Ghanbarian, B., 2024. A percolation model of unsaturated hydraulic conductivity using three-parameter Weibull distribution. Adv. Water Resour., 104696 https://doi.org/10.1016/j.advwatres.2024.104696.
Tartakovsky, A.M., Marrero, C.O., Perdikaris, P., Tartakovsky, G.D., Barajas-Solano, D., 2020. Physics-Informed deep neural networks for learning parameters and constitutive relationships in subsurface flow problems. Water Resour. Res. 56, e2019WR026731. https://doi.org/10.1029/2019WR026731.
Tehrani, A., Liaghat, A., Delbaz, R., 2025. Parameter estimation of soil water retention and thermal conductivity curves using HYDRUS-1D and inverse solution. Eur. J. Soil Sci. 76, e70095. https://doi.org/10.1111/ejss.70095.
Tehrani, A., Ziaei, A.N., Naghedifar, S.M., 2023. Irrigation scheduling of walnut seedlings using HYDRUS-1D and taguchi optimization approach. J. Irrig. Drain. Eng. 149, 04022045. https://doi.org/10.1061/(ASCE)IR.1943-4774.0001735.
Tuller, M., Or, D., 2005a. Water retention and characteristic curve. In: Encyclopedia of Soils in the Environment. Elsevier, pp. 278–289. https://doi.org/10.1016/B0-12-348530-4/00376-3.
Tuller, M., Or, D., 2005b. Water films and scaling of soil characteristic curves at low water contents. Water Resour. Res. 41. https://doi.org/10.1029/2005WR004142.
Turek, M.E., Pullens, J.W.M., Meurer, K.H.E., Moura Lima, E., Mehdi-Schulz, B., Holzkämper, A., 2025. Pedotransfer functions versus model structure: what drives variance in agro-hydrological model results? Eur. J. Soil Sci. 76, e70088. https://doi.org/10.1111/ejss.70088.
Van Genuchten, M.T., 1982. Analytical solutions of the one-dimensional convective-dispersive solute transport equation. US Department of Agriculture, Agricultural Research Service.
van Genuchten, M.T., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J. 44, 892–898. https://doi.org/10.2136/sssaj1980.03615995004400050002x.
Weber, T.K.D., Durner, W., Streck, T., Diamantopoulos, E., 2019. A modular framework for modeling unsaturated soil hydraulic properties over the full moisture range. Water Resour. Res. 55, 4994–5011. https://doi.org/10.1029/2018WR024584.