Pedotransfer functions in Earth system science: challenges and perspectives

Article

Published Version

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Pedotransfer Functions in Earth System Science: Challenges and Perspectives

Kris Van Looy[1,2] (iD), Johan Bouma[3] (iD), Michael Herbst[1] (iD), John Koestel[4] (iD), Budiman Minasny[5], Umakant Mishra[6] (iD), Carsten Montzka[1] (iD), Attila Nemes[7] (iD), Yakov A. Pachepsky[8] (iD), José Padarian[5] (iD), Marcel G. Schaap[9], Brigitta Tóth[10,11] (iD), Anne Verhoef[12] (iD), Jan Vanderborght[1] (iD), Martine J. van der Ploeg[13], Lutz Weihermüller[1], Steffen Zacharias[14] (iD), Yonggen Zhang[9,15] (iD), and Harry Vereecken[1,16] (iD)

[1]Institute of Bio- and Geosciences-3 Agrosphere, Forschungszentrum Jülich, Jülich, Germany, [2]Scientific Coordination Office ISMC, International Soil Modeling Consortium, Jülich, Germany, [3]Wageningen University and Research, Wageningen, Netherlands, [4]Swedish University of Agricultural Sciences, Uppsala, Sweden, [5]Department of Environmental Sciences, University of Sydney, Sydney, New South Wales, Australia, [6]Environmental Science Division, Argonne National Laboratory, Argonne, IL, USA, [7]Division of Environment and Natural Resources, Norwegian Institute of Bioeconomy Research, Ås, Norway, [8]Environmental Microbial and Food Safety Laboratory, USDA ARS Beltsville Agricultural Research Center, Beltsville, MD, USA, [9]Department of Soil, Water and Environmental Science, University of Arizona, Tucson, AZ, USA, [10]Institute for Soil Science and Agricultural Chemistry, Centre for Agricultural Research, Hungarian Academy of Sciences, Budapest, Hungary, [11]University of Pannonia, Georgikon Faculty, Department of Crop Production and Soil Science, Keszthely, Hungary, [12]Department of Geography and Environmental Science, University of Reading, Reading, UK, [13]Soil Physics and Land Management Group, Wageningen University and Research, Wageningen, Netherlands, [14]Department of Monitoring and Exploration Technologies, UFZ-Helmholtz Centre for Environmental Research, Leipzig, Germany, [15]Institute of Surface-Earth System Science, Tianjin University, Tianjin, China, [16]Centre for High-Performance Scientific Computing in Terrestrial Systems, HPSC TerrSys, Geoverbund ABC/J, Forschungszentrum Jülich GmbH, Jülich, Germany

Abstract Soil, through its various functions, plays a vital role in the Earth's ecosystems and provides multiple ecosystem services to humanity. Pedotransfer functions (PTFs) are simple to complex knowledge rules that relate available soil information to soil properties and variables that are needed to parameterize soil processes. In this paper, we review the existing PTFs and document the new generation of PTFs developed in the different disciplines of Earth system science. To meet the methodological challenges for a successful application in Earth system modeling, we emphasize that PTF development has to go hand in hand with suitable extrapolation and upscaling techniques such that the PTFs correctly represent the spatial heterogeneity of soils. PTFs should encompass the variability of the estimated soil property or process, in such a way that the estimation of parameters allows for validation and can also confidently provide for extrapolation and upscaling purposes capturing the spatial variation in soils. Most actively pursued recent developments are related to parameterizations of solute transport, heat exchange, soil respiration, and organic carbon content, root density, and vegetation water uptake. Further challenges are to be addressed in parameterization of soil erosivity and land use change impacts at multiple scales. We argue that a comprehensive set of PTFs can be applied throughout a wide range of disciplines of Earth system science, with emphasis on land surface models. Novel sensing techniques provide a true breakthrough for this, yet further improvements are necessary for methods to deal with uncertainty and to validate applications at global scale.

Plain Language Summary For the application of pedotransfer functions in current Earth system models, and specifically for the different fluxes of water, solutes, and gas between soil and atmosphere, subject of the land surface models, recent developments of knowledge are entered in a new generation of pedotransfer functions. Methods for development and evaluation of pedotransfer functions are described in this comprehensive review, and perspectives for future developments in different Earth system science disciplines are presented. Challenges are still present for the application in some extreme environments of the Earth. We argue that a comprehensive set of pedotransfer functions can be applied throughout a wide range of disciplines of Earth system science, with emphasis on land surface models. Even though methodological challenges are still present for extrapolation and scaling, as outlined, integration and validation in global-scale models is an achievable goal.

1. Introduction

1.1. Outline

An accurate description and prediction of soil processes and properties is essential in understanding the Earth system and the impacts of climate and land use changes. This, however, requires an accurate parameterization of soil processes and appropriate and reliable ways to represent the spatial heterogeneity of land surface. In many cases observations of essential soil properties, states and parameters that control water, energy, and gas fluxes of the terrestrial systems are not available due to the unfeasibility of performing measurements with sufficient spatial and temporal coverage. For the soil, hydrological, land surface, and climate sciences, obtaining accurate estimates of, for example, soil moisture and soil temperature, is essential for reducing the uncertainty in predicting soil respiration and heat and water fluxes. To this purpose, the soil science community has developed pedotransfer functions to estimate soil properties from data that are available from soil surveys (Figure 1).

Land surface models (LSMs) are a component of Earth system models (ESMs), which are key tools to predict the dynamics of land surface under changing climate and land use. However, reliable model representations of many critical variables (soil mineral content, water content, temperature, and carbon stocks) and processes (solute transport, heat, and water flow) of the soil system are needed in order to reduce the existing uncertainty in projecting the response of the terrestrial compartment of the Earth system (Beer et al., 2010; Friedlingstein et al., 2014; Qiu, 2014) under changing climate and land use scenarios. Improvement of the representation of soil properties that crucially determine prediction of soil states and related processes could alleviate shortcomings of LSMs (Pitman, 2003; Sato et al., 2015), for example, in the context of short to medium range weather forecasts (e.g., persistent temperature biases) or in the prediction of the carbon balance of critical biomes such as the Amazonian rain forest.

PTFs also play an important role in quantifying and predicting ecosystem services of soils (Vereecken et al., 2016). Ecosystem services of soils include regulatory services such as carbon sequestration and provisional services such as food supply and water storage. PTFs are used to quantify soil parameters and processes needed to estimate the delivery of ecosystem services and to quantify degrading and supporting processes. These services and processes are closely connected to societal issues and are key to scientific underpinning of our planet functioning (Adhikari & Hartemink, 2016). A discussion on pedotransfer functions, soil modeling, and ecosystem services is presented in Vereecken et al. (2016).

In this context, there is a high demand for high-resolution soil parameter estimation, necessary for improving land surface representations and predictions. We believe that the potential of available PTFs has not fully been exploited and integrated into LSMs and other models in Earth system sciences but also ecosystem services provided by soils. In addition, development of new PTFs might help in improving the description of soil processes and their parameterization. Access to novel databases such as Global Soil Map products (www.globalsoilmap.net) offers new opportunities to develop and apply PTFs in Earth system science. The grid-based soil information (organic carbon content, pH, sand, silt and clay content, bulk density, cation exchange capacity, and depth to bedrock) is being generated at various spatial resolutions (for, e.g., 1 km × 1 km to 250 m × 250 m) (e.g., Arrouays et al., 2014; Batjes, 2012; Food and Agriculture Organization (FAO), 2012; Hengl et al., 2014, 2017; Sanchez et al., 2009).

However, grid data represent points while the landscape dimension is essential when running hydrological, agronomic, ecological, or climate models. Just introducing spatially nonstructured point data in the development of PTFs, and implicitly assuming only one-dimensional (vertical) movement of water and nutrients, is not an appropriate assumption (van Tol et al., 2013). Information about small-scale spatial heterogeneity of soil properties is crucial for construction of knowledge rules for estimation of soil parameters that are used in large-scale models. The higher-resolution information supports improved PTFs and novel methods for extrapolation and upscaling that we discuss in this contribution.

Here the focus is on derivation of PTFs for improving parameterization of Earth system models, for which we envisage two possible approaches to bring about this improvement. One is looking at models' coefficients that are currently set constant, for example, temperature sensitivity coefficient Q10 (measure of rate of change following 10°C temperature increase, see section 5.4) values that are kept constant globally for all soils. Here PTFs can help to provide spatially distributed values that honor the impact of soil properties
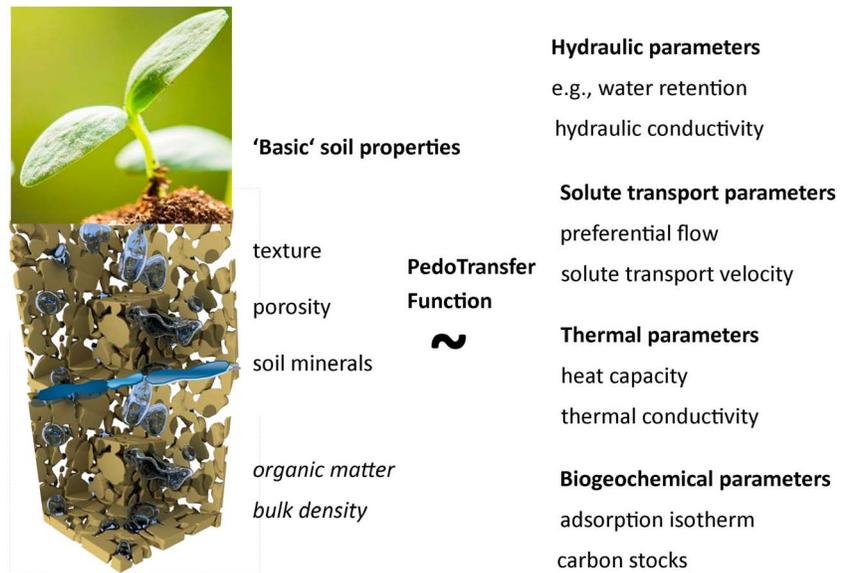
Figure 1. PTFs relate simple to measure soil properties to less available parameters of Earth system processes. Even for some more "standard" properties like soil organic matter and dry bulk density, PTFs are developed based on basic textural and structural properties trying to capture the biogeochemical processes context (this sums the four groups of parameterization dealt with: hydraulic, solute, thermal fluxes, and biogeochemical processes).

effecting Q10 values. The second avenue of improving parameterization is to translate new knowledge of environmental controls and variation of soil properties into spatially exploitable PTFs to parameterize specific processes. So either currently fixed parameters or proxies might be improved with PTFs, or equations and relationships that need parameterization could benefit from newly available methods for spatial extrapolations making use of the currently developed fine-grained soil information. In this paper we will show the opportunities for the Earth system modeling community to couple improved parameterization through novel PTFs with newly developed methods to spatial extrapolation and upscaling, integrating geographical and soil map information into LSMs with a suite of PTFs in combination with techniques like geographically weighted regression (Mishra & Riley, 2014, 2015).

The overall objective of this paper is to review the state of the art in developing and applying PTFs in Earth system sciences several reviews have been written in the past addressing specific topics and aspects of PTFs. McBratney et al. (2002) presented an overview of past and currently available PTFs with a list of physical, mechanical, and chemical properties estimated next to the classic soil hydraulic application and conclude that though most of the important soil physical properties are being predicted by PTFs, yet there is a striking absence for PTFs describing soil biological properties. Other reviews consider hydrological PTFs only (Vereecken et al., 2010; Wösten et al., 2001). Thus, there is not only scope for an update but a strong need for integration of existing knowledge on the linked processes in Earth system science with model developments trying to predict these processes in space and time for the cycles of water, energy, carbon, and nutrients. We specifically address the development and use of PTFs beyond the classical soil science and vadose zone research activities. In addition, we will highlight the use and development of PTFs in the framework of newly developed methods and databases that allow the coverage of large areas of estimated attributes derived from PTFs.

The paper is organized in the following sections. After a brief introduction of the historical context in section 1, we first refer to the methodological backgrounds and innovations in the derivation and evaluation of PTFs in section 2. Section 3 introduces the methodological challenges for PTFs in Earth system models. Sections 4 and 5 present the development and perspectives of PTFs for the vadose zone and land surface models, extending the long-standing experience on hydraulic characterization with current experience for solutes and thermal properties, onto novel applications for biogeochemical, biotic, and vegetation properties. Section 5 introduces the ways to new developments of PTFs by identifying weaknesses in parameterizations—either in variables entered as constants or in oversimplified functions—that can be solved by

improved process understanding and high-resolution data availability, illustrated for ecological and biogeo-chemical model elements, also in the context of climate change predictions. These confident perspectives to the applications of PTFs in Earth system models are continued in section 6 with an outlook to further improvements and a roadmap toward a full set of PTFs in Earth system modeling. The review ends with drawing conclusions in section 7.

1.2. A Brief History of PTFs

Estimation of soil properties from other more easily measurable soil properties has been a challenge in soil science from its early beginning. Earliest estimation equations date back to the beginning of twentieth century. They relate, by regression, for the first time the soil moisture characteristics to soil texture (Briggs & Lane, 1907; Veihmeyer & Hendrickson, 1927). By the second half of the century, these equations were well anchored in extensively used soil classification and cartography initiatives. The concept of *pedotransfer function* (PTF) proposed by Bouma (1989) basically embraces all these earlier estimation approaches, most of which have been reviewed by McBratney et al. (2001). Bouma (1989) described its concept as translating soil data we have into data we need but do not have, and its introduction was part of a framework of quantitative land evaluation. The proposal was initiated at a time when soil surveys in many countries were either completed or terminated and questions were raised as to what would be a logical next activity of pedology, the science of studying the formation, occurrence, and behavior of soils in the field (Bouma, 1988). In land evaluation, interpretations of soil surveys are made in terms of limitations for a wide range of uses and also in terms of suitability (Bouma et al., 2012; FAO, 2012). Already in 1989 it was clear that such qualitative and empirical judgements could be relevant when producing initial assessments over large areas. However, they would require more quantitative procedures to be adequate enough to face modern land use questions. The key aspect of the PTF proposal was the link between pedology and soil survey, allowing a comprehensive landscape perspective based on localized samples. More specifically, PTFs aim at transferring structural and compositional data of soils into information that characterizes soil functioning (and thus allows parame-terization), such as parameters to define soil hydraulic functions, mineralization constants, and sorption prop-erties. This information can then lead to quantification of soil ecosystem services such as providing water and nutrients to plants, regulating climate and biogeochemical cycles, buffering and filtering of solutes, and trafficability/accessibility of soils, for example, through the use of mathematical process models (Bouma, 1989; Vereecken et al., 2016). The data used in PTFs include soil horizonation, color patterns, texture, qualita-tive structural and morphological information, organic matter content, pH, redox, and mineral concentrations (Bouma, 1989; McBratney et al., 2002). Moreover, these data have spatial attributes and cover the land surface of countries and continents making them valuable for the use in simulation of processes occurring at the land surface. Bouma (1989) recognized that simulation models of soil processes such as water flow, heat flow, and solute transport at the field and catchment scale were already operational in the 1980s and model develop-ment since that time has been remarkable. These models need a number of parameters; most of which can, in principle, be measured but at great expense and effort. The proposed PTFs attempt, therefore, to link global data obtained in soil surveys, which are nowadays globally available in databases with high resolution, to model parameters.

Bouma (1989) distinguished two types of PTFs: continuous and class PTFs. Continuous PTFs use continuous quantities such as clay, sand, or organic matter content. Class pedotransfer functions relate modeling parameters to classes of soil properties, as distinguished in a soil surveys (Bouma, 1989). Baker and Bouma (1975) made, for example, multiple measurements of water retention and hydraulic conductivity curves in subsurface horizons of silt loam soils that were classified as Tama silt loam. Curves obtained formed relatively narrow bands, expressing spatial variability. When making new measurements in Tama silt loams at other locations, results obtained fitted well within the bands that were earlier established, demonstrating the possibility to extrapolate measured data in a given soil series to sites without measurements but with the same soil classification. The soil series is thus used as a class pedotransfer function that can also relate to particular soil horizons within a given soil series or to more general texture classes (Baker, 1978). Wösten et al. (1986) made a similar analysis for soil horizons in the Netherlands and compared four methods to generate soil hydraulic functions, including direct measurements and use of both class and continuous PTFs (Wösten et al., 1990). Differences between the methods were not statistically significant, demonstrating the potential of both types of PTF.

Currently, continuous PTFs are typically used to parametrize soil processes in simulation models of water, energy, and carbon cycles from the field to the continental scale. Continuous PTFs can in this respect be divided/classified as point or parametric PTFs (Vereecken et al., 2010). Point PTFs estimate, for example, specific points of the water retention curve such as wilting point (defined as the minimal point of soil moisture the plant requires not to wilt), whereas parametric PTFs estimate, for example, the parameters of the Mualem-van Genuchten (MvG) model (see section 4.1.1; van Genuchten, 1980). The continuous PTFs also capture the early work that was done on deriving soil hydraulic properties such as the water retention characteristic, unsaturated and saturated hydraulic conductivity, and soil water content at prescribed pressure heads for simple soil properties (Bloemen, 1977, 1980; Clapp & Hornberger, 1978; Cosby et al., 1984) and which has been reviewed by, for example, Wösten et al. (2001), Rawls and Pachepsky (2002a), and Vereecken et al. (2010). In fact, most of the early work on PTFs has focused on estimation of soil hydraulic parameters of the water retention characteristic and hydraulic conductivity functions (e.g., Vereecken et al., 2010; Wösten et al., 2001) as these parameters are difficult and cumbersome to measure but key for all simulations of water, matter, and energy fluxes of the land surface. Later on, PTFs were also developed for other soil properties or soil functions beyond soil water flow (Breeuwsma et al., 1986; McBratney et al., 2002). Some recent PTFs make distinctions between top soil and subsoil hydraulic parameters (Wösten et al., 1999) and include further predictors such as chemical soil properties ($CaCO_3$, pH, and cation exchange capacity (CEC)) to estimate hydraulic parameters (Botula et al., 2013; Pachepsky & Rawls, 1999; Tóth et al., 2015).

These hydraulic parameter estimations are not only the best known and most developed examples of PTFs; they also are at the basis of the construction of other PTFs—for example, of soil thermal conductivity (see section 4.3)—and exemplary to the development of other transfer functions. In this contribution, we propose to benefit from this exceptional richness in methodological development of parameter estimation and go beyond its classic application for hydraulic parameters.

2. Methods to Derive and Evaluate Pedotransfer Functions

2.1. General Considerations

Before we discuss the wide variety of methods available, it is worthwhile to discuss the general concept of PTF development as illustrated in Figure 2. The general purpose of PTF development is to establish predictive models using databases of soil properties which hold suitable predictors ("basic soil properties") and desired "estimands" ("estimated less available soil properties"). Such databases are often highly specialized because they must hold both predictor and estimand observation data, and as a rule, they are often much smaller (typically hundreds to thousands of records) than soil survey-based pedological databases (refer to National Cooperative Soil Survey, 2017, WoSIS (Batjes et al., 2017)), which hold tens or even hundreds of thousands of records. Databases used in PTF development may therefore not reflect the true population of soils in a region or the world, and as a result, PTFs tend to be biased to the database on which they were calibrated (Schaap & Leij, 1998a). Once a PTF is calibrated, it is usually published in the form of a relatively simple mathematical formula, a look-up table, or a software package. A typical "user" will then incorporate the PTF in their work, which usually requires the PTF outcome to be transformed in a relevant way. The user of a PTF may therefore be interested in other soil attributes (e.g., infiltration capacity) than for which the PTF was specifically calibrated (e.g., saturated hydraulic conductivity or key parameters of the moisture release curve).

Beyond some general conceptual understanding there are no precise a priori relations that link predictors with the estimands. In addition, most PTFs differ with regard to the set of predictors ("input variables") and with regard to the estimands ("output variables"). Typical sets of predictors are therefore sand, silt, and clay percentages (Campbell & Shiozawa, 1992; Cosby et al., 1984; Rawls & Brakensiek, 1985; Saxton et al., 1986), bulk density and organic carbon, or organic matter content (De Lannoy et al., 2014; Rawles & Brakensiek, 1982; Rawls et al., 1983; Saxton & Rawls, 2006; Vereecken et al., 1989; Weynants et al., 2009; Wösten et al., 1999)—though morphological properties (Ali & Biswas, 1968; Lin et al., 1999) and soil structural information (Nguyen et al., 2015; Pachepsky & Rawls, 2004; Pachepsky et al., 2006; Rawls & Pachepsky, 2002b) have also been used. Some PTFs even use one or several moisture retention points to improve their estimation of the rest of the water retention curve (Ahuja et al., 1985; Paydar & Cresswell, 1996; Rawles & Brakensiek, 1982; Schaap et al., 2001; Zhang & Schaap, 2017). Finally, recognizing that users are often faced with different
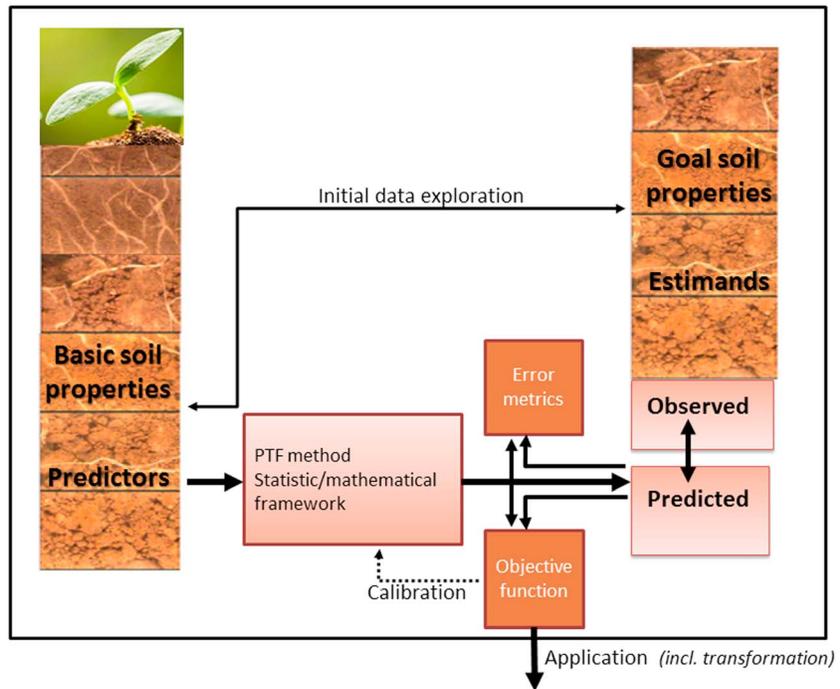
Figure 2. General concept of PTF development, based on a calibration database with both the basic and "estimand" soil properties measured.

levels of data availability, some researchers developed PTFs that can deal with limited and more extended sets of predictors (Nemes et al., 2003; Schaap et al., 2001, 2004; Tóth et al., 2015; Twarakavi et al., 2009; Zhang & Schaap, 2017).

The mathematical/statistical frameworks to establish PTFs, that is, the derived relationships between predictors and estimands (Figure 2), have been extremely diverse. While some (semi)empirical approaches exist, most methods are entirely empirical, purely data driven and vary from look-up tables that provide soil hydraulic parameters for specific soil textural classes to simple linear/nonlinear regression models, to more sophisticated data mining methods such as artificial neural networks (ANNs), regression and classification trees and their derivatives (e.g., classification and regression trees, chi-square automatic interaction detection (CHAID), boosted trees, and random forests), k-nearest neighbor-type algorithms and support vector machines (SVM), and some other methods that are less commonly used, like, for example, group method of handling data (Nemes et al., 2005; Pachepsky & Rawls, 1999).

2.2. Look-Up Tables and Class PTFs

The simplest PTFs, yet widely applied, are the look-up tables that provide textural class-average hydraulic parameters (Baker, 1978; Bouma, 1989). Such look-up table is the first step to identify the dependence of soil hydraulic parameters on soil texture class (Cosby et al., 1984). Because of simplicity, the look-up table has been widely used in soil sciences and other disciplines, such as land surface modeling. For example, Rosetta model H1 (Schaap et al., 2001), which is essentially a look-up table, is incorporated in variably saturated media codes Hydrus 1-D, 2-D, and 3-D. The look-up table developed by Cosby et al. (1984) is widely used in land surface modeling, such as the Biosphere-Atmosphere Transfer Scheme developed by Dickinson et al. (1986, 1993) and the Global Land Data Assimilation System developed by Rodell et al. (2004) to estimate soil hydraulic parameters. Today, look-up tables are also included in several widely used soil survey guidelines, with textural class averages for, for example, field capacity, values of saturated hydraulic conductivity, and even pressure-based descriptions of "typical" water retention curves. The drawback of look-up table PTFs is the variability of many parameters within soil textural classes. For example, depending on the measurement, very different soil water retention parameters and saturated hydraulic conductivity (K_s) can be documented per texture class. Table 1 lists K_s values calculated for U.S. Department of Agriculture (USDA) soil textural classes in selected publications. Cosby et al. (1984) and Carsel and Parrish (1988) have

Table 1

K_s Values of USDA Soil Texture Classes in Look-Up Tables of Selected Publications

Texture class	Cosby et al. (1984)	Carsel and Parrish (1988)	Zhang and Schaap (2017)
Clay	8.4	4.8	14.8
Silty clay	11.6	0.5	9.6
Sandy clay	62.4	2.9	11.4
Clay loam	21.1	6.2	7.1
Silty clay loam	17.6	1.78	11.1
Sandy clay loam	38.5	-	13.2
Loam	29.2	25.0	13.3
Silt loam	24.3	10.8	18.5
Sandy loam	45.2	106.1	37.5
Silt	-	6.0	43.8
Loamy sand	121.7	350.2	108.2
Sand	402.8	712.8	643.0

Note. All results are transformed to the same unit: cm/d.

data from the USA, with 1,448 and 5,097 data sets, respectively, while Zhang and Schaap (2017) used 1,306 data sets from the USA and Europe. Carsel and Parrish (1988) estimated lower K_s values for fine soil textural classes, but higher K_s values for coarse soil texture compared with the other two soil PTFs. Cosby et al. (1984) estimated higher K_s values compared with the estimations of Zhang and Schaap (2017) with some exceptions. Such important differences between PTFs may result from the calibration data and from different methods to develop PTFs. It immediately stresses the need for contextualization (see extrapolation and upscaling in section 3) of the application of PTFs.

Recently, Twarakavi et al. (2010) observed the validity of texture-based classifications for soil hydraulic studies and compared it against a classification based on soil hydraulic characteristics. Although they found similarities between both classification schemes, they found larger differences for soils with lower sand contents where the water flow was dominated by capillary forces. However, the soil hydraulic classification led to only marginal improvements regarding the prediction of soil hydraulic parameters compared to the classical soil textural classification. The authors attributed the lack of comprehensive soil hydraulic databases as main bottleneck for the further development of soil classification systems.

2.3. Regression Techniques

Regression technique is widely used to determine the relationship between predictors and estimands because of its simplicity. Regression analysis can use linear regressions or nonlinear regressions depending on the expected relationship among variables. They typically follow the general form:

$$p = a^* \text{Sand} + b^* \text{Silt} + c^* \text{dry bulk density} + \cdots + x^* \text{var} X \qquad (1)$$

where p is the soil property that is to be estimated. a, b, c, and x are regression coefficients, and var$X$ any other basic soil property that can easily be measured (Wösten et al., 2001). Gupta and Larson (1979) were probably among the first to build a linear regression to estimate soil water retention characteristics from soil particle size distribution, organic matter, and bulk density. Rawls and Brakensiek (1985) built an exponential relationship between soil hydraulic parameters and soil texture as well as bulk density. The advantage of regression analysis is that it is straightforward to carry out and easy to be used, while the disadvantage is that the regression equations (for example, linear, logarithmic, or exponential) and predictors have to be determined as a priori and that the relationships between soil properties and predictors may be different in different portions of the database. Boosted multiple linear regression can be an efficient method if relationship between dependent and independent variables is not complex. Cosby et al. (1984) performed analysis of variance to determine the significance of hydraulic parameters with the predictors and then established relationships to estimate hydraulic parameters based on univariate and multivariate functions of the predictors.

2.4. Neural Networks

Artificial neural networks (ANNs) require no a priori model concept. Being described as universal function approximators, ANNs are the powerful technology available modeling complex "input-output" relationships (Haykin, 1994; Hecht-Nielsen, 1990; Maren et al., 2014). The relationship can then be used similar to a regression formula to make predictions of soil properties.

Pachepsky et al. (1996), Schaap and Bouten (1996), and Tamari et al. (1996) are among the first to build PTFs to estimate soil hydraulic parameters using ANNs, and later publications that further pursued this topic include Schaap and Leij (1998b), Schaap et al. (1998, 2004), Pachepsky and Rawls (1999), Minasny and McBratney (2002), Nemes et al. (2003), Minasny et al. (2004), Sharma et al. (2006), Parasuraman et al. (2006), Agyare et al. (2007), Ye et al. (2007), Baker and Ellison (2008a), Jana and Mohanty (2011), Haghverdi et al. (2012, 2014), and Zhang and Schaap (2017). It should be stressed that there are many types of neural networks (Hecht-Nielsen, 1990). Among these, so-called feed-forward backpropagation ANNs are most widely used to map input-output relationships. A typical feed-forward backpropagation ANN contains input, hidden, and output neurons. The hidden neurons extract useful information from

the input and utilize them to predict the output, with the determination of the number of hidden neurons empirically or based on the performance of calibration and validation data set (Zhang & Schaap, 2017). The input vector of neurons x_j ($j = 1...J$) in network is weighted, summed, and biased to produce the hidden neurons y_k ($k = 1...K$):

$$y_k = \sum_{j=1}^{J} w_{jk}x_j + b_k, \qquad (2)$$

where J is the number of input neurons and k is the number of hidden neurons. The hidden neurons consist of the weighted (w_{jk}) input and a bias (b_k). The hidden neurons y_k are then operated by an activation or transfer function f to produce

$$r_k = f(y_k). \qquad (3)$$

The activation function is usually a monotonic function, which can reflect the nonlinearity in the input-output relationship. Commonly used activation functions include sigmoid, hyperbolic, tansig, and pure linear functions.

The output from the hidden neurons is processed by a similar procedure to that in equation (2) as follows:

$$v_l = \sum_{k=1}^{K} u_{kl}r_k + b_l \qquad (4)$$

and then are transformed by another activation function F to produce the output z:

$$z_l = F(v_l). \qquad (5)$$

The weights and biases are obtained in ANN by minimizing the following objective function through an iterative procedure,

$$O(w_{jk}, b_k, u_{kl}, b_l) = \sum_{n=1}^{N_s} \sum_{m=1}^{N_p} \left[t_{n,m} - t'_{n,m}(w_{jk}, b_k, u_{kl}, b_l) \right]^2 \qquad (6)$$

where N_s is the number of calibration samples, N_p the number of parameters, and t and t' the observed and predicted variables (see Figure 2). Originally, the backpropagation algorithm was used to minimize the above objective function, while other alternative algorithms such as Levenberg-Marquardt are also available (Press et al., 1992).

Although several studies suggest that ANNs perform better than regression-based PTFs (Schaap & Bouten, 1996; Schaap et al., 1998; Tamari et al., 1996), there are still several issues with the ANN use. One of the problems is overfitting, which means that, as the fitting proceeds, the objective function (equation (6)) will continue to decrease for the calibration data set, while the objective function will eventually start to rise for the validation data set. The implementation of cross-validation methods (Hastie et al., 2001) is an approach to help overcome the issue by dividing the data set into calibration and validation, with the most optimal ANN models being where the validation error is the lowest (Schaap et al., 2001, 2004; Zhang & Schaap, 2017). Another drawback of ANNs is that they often contain a considerable number of coefficients, which prevents publication of the PTF as a closed-form equation, especially when combined with the bootstrap method (Efron & Tibshirani, 1993).

2.5. Support Vector Algorithms

Support vector machines (SVMs) are another data mining tool frequently applied to build PTFs. SVMs use a supervised nonparametric statistical learning method, which is originally presented with a set of labeled data, and SVM training is to find a hyperplane which can separate the data set into a discrete predefined number of classes (Vapnik, 2013; Vapnik & Vapnik, 1998). The optimal separation hyperplane is a decision boundary that minimizes misclassifications, obtained in the calibration process in an iterative way (Mountrakis et al., 2011; Zhu & Blumberg, 2002). In their simplest form, SVMs are linear binary classifiers that allocate a class (from one of the possible labels) to a given test sample. In practice, linear separability is not often present.To divide the data set into classes accurately in such cases, the use of kernel functions can solve this problem.

Suppose that we have training data set $(x_i, y_i)(i = 1, \ldots, n_t)$, where x_i is the data point, y_i corresponding properties of data point x_i. In $\varepsilon-$ support vector regression, the goal is to find a function $f(x)$ to estimate an unknown variable x that has at most ε deviation from the actually obtained targets y_i for all the training data, and that is as flat as possible.

The regression is formulated as

$$f(x) = \sum_{i=1}^{l} w_i \Phi(x) + b, \tag{7}$$

where $\Phi(x)$ denotes nonlinear transformation (using kernel functions) that map data into better representation space, making nonseparable problem separable. w_i are weights and b the bias term; both w_i and b are parameters estimated by minimizing the following objective function:

$$\frac{1}{2}|w|^2 + C \sum_{i=1}^{l} (\xi + \xi^*) \tag{8}$$

subject to

$$\begin{cases} y_i - f(x) - b \leq \varepsilon + \xi_i^* \\ f(x) + b - y_i \leq \varepsilon + \xi_i, \\ \xi_i, \xi_i^* \geq 0, i = 1, \ldots l \end{cases} \tag{9}$$

where $|w|^2 = \sum_{i=1}^{l} w_i^2 . \xi, \xi^*$ are slack variables introduced to cope with infeasible constrains (Vapnik, 2013).

The cost parameter $C > 0$ determines the trade-off between the complexity of the SVM structure and the amount up to which deviations larger than ε are tolerated. The insensitivity parameter, ε, controls the width of the insensitive zone; larger values of ε will lead to smaller numbers of support vectors and result in poor generalization (Twarakavi et al., 2009). A number of kernel function are available (Vapnik, 2013; Vapnik & Vapnik, 1998). Twarakavi et al. (2009) used radial basis kernel function to perform the nonlinear transformation.

Lamorski et al. (2008) developed PTFs with a soil data set in Poland using SVMs and ANNs. They found that SVMs outperformed or had the same accuracy compared with ANNs. Twarakavi et al. (2009) developed SVM-based PTFs using the data set of Rosetta (Schaap et al., 2001) and found SVMs outperforming ANN-based PTFs. More recent applications include Skalová et al. (2011), Haghverdi et al. (2014), Nguyen et al. (2015), and Khlosi et al. (2016).

2.6. *k*-Nearest Neighbor Methods

The k-nearest neighbors (KNN) method is another machine learning algorithm, which has been used to derive PTFs for soil properties (Jagtap et al., 2004; Nemes, Rawls, & Pachepsky, 2006; Nemes, Rawls, Pachepsky, & Van Genuchten, 2006). KNN uses a distance-based approach where the distance for each soil from a target soil can be calculated as the square root of the sum of squared differences in predictors between the target soil and each of the soils of a reference data set that corresponds to the development data set in other techniques. As the units of the predictors can be different, for example, sand fraction between 0 and 100, but organic matter content between 0 and a maximum of 15% in nonorganic soils, a potential bias to one or another variable is avoided either by prior normalization (e.g., Nemes, Rawls, Pachepsky, & Van Genuchten, 2006) or by using an extended standardized Euclidean distance metric that takes into account the covariance structure among the predictor variables (cf. Mahalanobis distance; e.g., Tranter et al., 2009). The output of KNN is the property value for the object, and the value is the average of soil property values of its k-nearest neighbors for the estimated soil samples. The optimal choice of k depends upon the data, and the user can choose among different weighting schemes of the selected samples. In a simple case, the resemblance of each sample in the reference data to the sample in question is evaluated by calculating its "distance" in terms of the previously normalized variables as

$$d_i = \sqrt{\sum_{j=1}^{x} \Delta a_{ij}^2} \tag{10}$$

where d_i is the "distance" of the ith soil from the target soil and Δa_{ij} is the difference of the ith soil from the target soil in the jth soil attribute. The user then sets how many samples (k) are necessary to account for in formulating the estimate, and a weight term is then assigned to each of the selected k soils. As opposed to simple averaging or rank-dependent weighting solutions found in literature (e.g., Lall & Sharma, 1996), Nemes, Rawls, and Pachepsky (2006) proposed a distance-dependent weighting as

$$w_i = \left(\sum_{i=1}^{k} d_i \ / \ d_i \right)^p \ / \ \sum_{i=1}^{k} \left(\sum_{i=1}^{k} d_i \ / \ d_i \right)^p, \tag{11}$$

where k is the number of nearest neighbors considered, w_i is the assigned weight, d_i is the "distance" value of the ith selected neighbor calculated as above, and p is a power term, which was found to be dependent on the size of the reference data set (Nemes, Rawls, & Pachepsky, 2006). The final output is then formulated as the weighted sum of observed values of the output variable of the selected k-nearest neighbors.

Nemes, Rawls, and Pachepsky (2006) developed PTFs using 2,125 soil samples from U.S. normalized radar cross section soil characterization database by utilizing ANN and KNN. They found that the KNN method has a nearly identical performance to that of ANN in terms of the evaluated criteria. Nemes, Rawls, Pachepsky, and Van Genuchten (2006) analyzed the sensitivity of KNN variant to different data and algorithms. Subsequent applications of KNN to develop PTFs include, for example, Ghehi et al. (2012), Botula et al. (2013), and Nguyen et al. (2015). A kriging-based Gaussian process approach similarly utilizes nearest neighbors, measuring the "distance" between neighbors based on a covariance function (Rasmussen, 2004).

2.7. Decision/Regression Trees

Different parts of the data set may have different PTF dependencies, and using one unique PTF equation for the entire data set may not be justified. It may be beneficial to split the entire data set into homogeneous parts and develop independent PTFs for different parts of the data set (Schaap, 2004).

Decision trees are recursive data partitioning algorithms that have a continuous response variable and recurringly subdivide the presented data into two subsets, making subsets as homogeneous as possible at each level of partitioning (Breiman et al., 1984). Each partitioning can be viewed as a branching of tree (Wösten et al., 2001). Regression trees have continuous response variable, which is the case most frequently in PTFs. For categorical-type dependent variable classification trees can be applied. In regression-type problems the partitioning is to divide the data into $R_1, \ldots, R_J$ subsets by minimizing the residual sum of squares (RSS) (equation (12)):

$$\text{RSS} = \sum_{j=1}^{J} \sum_{i=1}^{n_{Rj}} \left(y_i - \hat{y}_{Rj} \right)^2 \tag{12}$$

where J is the number of subsets, n_{Rj} is the number of observations belonging to R_j subset, $\hat{y}_{Rj}$ is the mean response for the training observations within the jth set, and y_i is the observation in R_j subset.

The most important features of recursive partitioning methods are reviewed in Strobl et al. (2009). Another possibility for recursive partitioning is the chi-square automatic interaction detector (CHAID) method (Kass, 1980) in which the stopping criterion is based on statistical significance tests. The independent variable with highest association to the dependent variable is selected for splitting. Efficiency of CHAID was found to be similar to that of regression trees for the derivation of point PTFs on Hungarian salt-affected soils (Tóth et al., 2012).

The regression tree approach was first used to develop PTFs by McKenzie and Jacquier (1997). Subsequent applications of regression trees to develop PTFs include McKenzie and Ryan (1999), Rawls and Pachepsky (2002a), Pachepsky and Rawls (2003), Pachepsky et al. (2006), Lilly et al. (2008), Nemes et al. (2011), Gharahi Ghehi et al. (2012), Koestel and Jorda (2014), Jorda et al. (2015), and Tóth et al. (2015).

A limitation of the regression tree is that it only produces an estimate under a "terminal node," which can cause a discontinuity in the response function (Staver & Hansen, 2015). Another advanced form of regression tree is called the model tree or M5 or Cubist model (Solomatine & Dulal, 2003), which produces a linear regression at the terminal node. Recent progress includes using an ensemble of many regression trees in a procedure called boosting which was used to identify qualitative/categorical soil properties that help

improve the estimation of K_s (Lilly et al., 2008) and derive a PTF of bulk density for forest soils (Jalabert et al., 2010) or K_s (Jorda et al., 2015).

2.8. Random Forest

A single regression tree is limited in accuracy, to solve this problem, and therefore, an ensemble of trees has been introduced (Breiman, 2001). During the construction of the model lots of regression trees are grown with randomly selected combination of input variables. In this way the model will be more robust to outliers and noise than a single regression tree. Prediction is based on a whole set of regression trees, while the results of all individual trees are averaged or weighted average is calculated.

This technique has been widely used in recent years (Akpa et al., 2016; Koestel & Jorda, 2014; Sequeira et al., 2014) given its reported prediction performance. De Souza et al. (2016) analyzed the performance of the random forest technique in predicting bulk density based on soil properties and environmental covariates. Koestel and Jorda (2014) used this method to predict the strength of preferential transport. Tóth et al. (2014) developed point PTFs to predict water retention at 0, −33, −1,500, and −150,000 kPa matric potential with the conditional inference forest (cforest) method, a random forest method based on conditional inference trees (ctree) (Strobl et al., 2008). The advantage of the method is that selection of variables is unbiased when independent variables measured at different scales (e.g., clay content—interval, soil type—nominal, and topsoil and subsoil distinction—dichotomous) are used—that is, in ctree continuous variables or predictors with more categories are not favored—because statistics analyzing the relationship between independent and dependent variables use conditional distribution of those (Hothorn et al., 2006). As it is sometimes being used as a silver bullet, we advise to use with caution due to its "unexpected" behavior when dealing with noisy data (Segal, 2004).

2.9. PTF Evaluation Criteria
2.9.1. Evaluation Methods

Donatelli et al. (2004) and Schaap (2004) reviewed various methods to evaluate and quantify the quality of PTFs for predicting soil water retention parameters and hydraulic conductivities. These methods are applicable to any type of PTF developed. The most common metrics used to evaluate PTF performance are root-mean-square errors (RMSEs), mean errors (MEs), and the coefficient of determination (R^2). RMSE values quantify the root of the average bivariate variance between estimated and measured quantities. ME quantifies systematic errors or bias. Negative ME values indicate an average underestimation of the quantity being evaluated, while positive values indicate an overestimation of target variables. For a truly well performing PTF, both RMSE and ME should be as low as possible. We note that the ME pertains to an average estimation over N data points. So it is possible that ME is 0, but the PTF, for example, overestimates soil properties for coarse-texture soils and underestimates soil properties for fine-texture soils. In case the overestimation and underestimation might cancel out, absolute mean errors can be computed.

Sometimes, it is useful to quantify the relative size of the systematic errors, which can be computed using the relative mean errors (RME), or the unbiased RMSE (URMSE) values, succesfully used by Tietje and Hennings (1996) and Schaap et al. (2004) that separate the systematic errors from the random errors. URMSE values should always be equal to or smaller than the corresponding RMSE value.

For evaluating PTF performance and the development method, both the choice of the data set (local-regional, within or across soil types) and the range of input variables play crucial roles. Results of such evaluations to the methodological performance need to be considered with care. For example, many techniques were tested to estimate bulk density based on clay and sand content, calcium carbonate equivalent, pH, and organic carbon content, with differing conclusions to the best performing technique of PTF development (Table 2). Shiri et al. (2017) documented that for this application the novel gene expression programming technique performed strongest, but the differences from the other methods were not significant.

Typically not as a metric to be published with a PTF, but as a diagnostic tool while deriving PTFs, one should evaluate patterns in the estimation residuals and correlations between residuals and input or output properties (Boschi et al., 2014, 2015). This has helped, for instance, to diagnose and improve models (e.g., Nemes, Rawls, Pachepsky, & Van Genuchten, 2006; Nemes et al., 2011) or to help understand sources of errors and differences between models (Nemes et al., 2009).

Table 2
Examples of Documented Bulk Density PTF Development Method Evaluation and Comparison

Bulk density PTF method evaluation	Linear regression	Nonlinear regression	ANNs	SVMs	KNN	Regression trees	Random forest
Jalabert et al. (2010)						X	
Ghehi et al. (2012)					X	X	
Patil and Chaturvedi (2012)	X		XXX				
Al-Qinna and Jaber (2013)	X	X	XXX				
Botula et al. (2015)	X				X		
Rodríguez-Lado et al. (2015)			X				XXX
Xiangsheng, Guosheng, and Yanyu (2016)	X		XXX				
Shiri et al. (2017)			X	X		X	X

Note. X successfully applied method and XXX strongest performing method in evaluation.

It is plausible to analyze and compare the usefulness of PTFs by using them as input information in Earth system models and evaluating the Earth system model performance rather than just PTF performance. Such "functional evaluation" can clearly quantify the utility and value of the PTFs (Vereecken et al., 1992; Xevi et al., 1997). Finally, PTFs form no goal in themselves; their function is in estimating functional soil properties that users are interested in such as water supply capacity and leaching of chemicals. It may very well turn out that calculated properties using PTFs are different from measured ones but that when put into models as parameters, modeling results may not significantly differ. The differences between measured and calculated parameters should certainly be established, but the value of the evaluation process is increased when the next step is taken as well. Especially, when an application is oriented to complex models at continental to global scales, there can be a trade-off between gain in precision and accuracy of prediction, which may be triggered (and optimized) by the method of PTF development as well.

2.9.2. Strengths and Weaknesses of PTF Development Methods

When selecting PTF development techniques for future work, many authors focus purely on reported model performance in terms of the above metrics. It is, however, recommended that other characteristics of modeling techniques are considered as well, especially since the differences in reported performance are often very small and/or inconsistent, and their significance is reduced when functionally tested in an application. We provide some guidance on some strengths and weaknesses of popular PTF development techniques cited in this paper (Table 3). Note that this is a rough, generic guide that cannot account for the variety of all software options that one has—and will have—to implement a particular technique. Software keeps evolving, and a modeler always has implementation options whose features may differ from those of off-the-market software.

Table 3 presents a nonexhaustive list of seven categories of model qualities that may be useful for a user to consult prior to choosing modeling techniques. It is apparent that strengths and weaknesses often come as trade-offs, although that is not a rule. Users may want to capitalize on strengths that are essential for the purpose of their study. For example, PTF studies may assume one of two primary purposes: research or application. Authors who intend to advance research knowledge may find it more desirable that the model is flexible and can work with various data sizes and types efficiently or help mine auxiliary information (e.g., variable importance) given their structure or features. Application-oriented

Table 3
Comparison of Different Mathematical Predictive Models, ++ = Good, + = Fair, and − = Poor (Adapted From Hastie et al., 2001)

Feature	Class PTF	MLR, GLM	GAM	Regression tree	Random forests	Neural net	SVM	Nearest neighbor
Parsimony	++	++	−	++	−	−	+	−
Interpretability of the model	++	++	+	++	−	−	−	−
Variable selection	−	++	−	++	++	−	−	−
Nonlinearity	−	−	++	++	++	++	++	++
Handling of mixed data type (qualitative and quantitative)	+	+	+	++	++	−	++	+
Computational efficiency (large data)	++	++	+	++	−	−	+	++
Predictive power	−	+	+	+	++	++	++	++
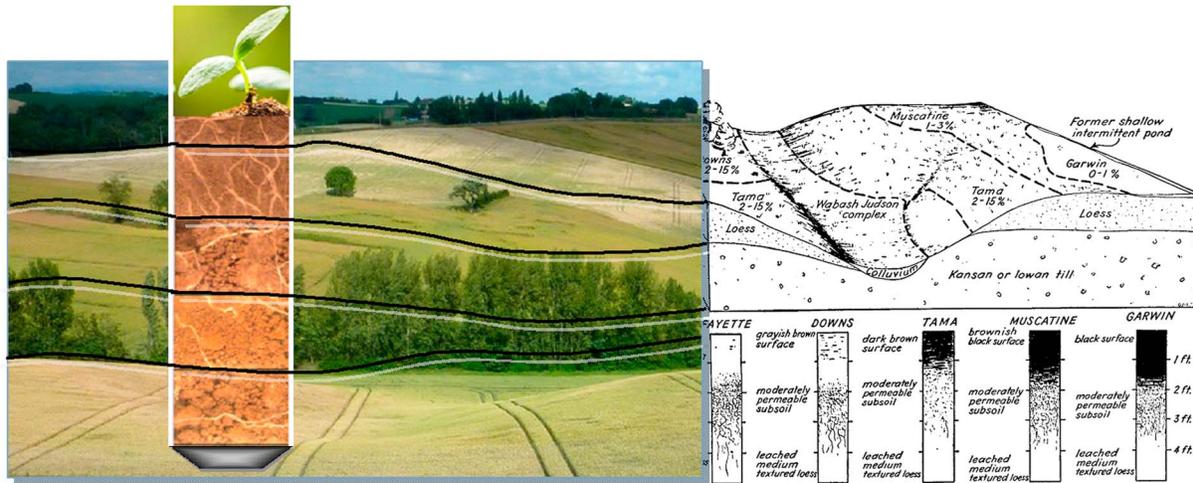
Figure 3. Simplified illustration of the interpretation of landscape topography and soilmaps (background USDA-SCS diagram) to soil structure from soil samples.

studies may benefit from better transparency, interpretability, and ease of applicability. There is also an increasing need for application oriented PTF studies that are designed to consider inputs or input levels that serve large-scale applications, as was recently done with the hydraulic pedotransfer functions for Europe (EU-HYDI PTFs) (Tóth et al., 2015).

3. Methodological Challenges for PTF Use in Earth System Sciences

Theoretical understanding of soil formation suggests that soil properties can be predicted as a function of soil-forming factors such as climate, biota, topography, parent material, and time (Jenny, 1941; McBratney et al., 2002). Information of these soil-forming factors can be used to capture and predict the spatial variation of soil properties, for example, illustrated for soil organic carbon (SOC) stocks over Alaska (Mishra & Riley, 2012; Vitharana et al., 2017). The high-resolution soil information available nowadays allows for improved PTFs and improved methods for extrapolation and upscaling. PTFs are basic tools to extrapolate knowledge on soil properties from one location to another, for their application over larger geographical entities (region to global) constrained by the understanding of interactions with the soil-forming factors. Furthermore, they are critically determined by scale; PTFs derived in the lab at the pore scale are not applicable at field scale, and estimations at the landscape level do not comply with regional-scale estimates. Sources of information on soil variability are essential for the application of PTFs. Topographical and geographical information, together with soil maps, can contribute to extrapolation of soil properties like layer depth, structure, compaction, and organic content (Figure 3).

In this perspective of development of PTFs for land surface models, the spatial interpretation of actual high-resolution soil information (e.g., SoilGrids https://soilgrids.org) is essential for estimating relations between soil properties and needs to be combined with topographical and geographical information in extrapolation and upscaling. Digital terrain models can provide detailed information on surface topography and that, in combination with knowledge about soil types and soil properties in a landscape, can improve simulation of 3-D landscape processes. This is a crucial step in optimizing the applicability of PTFs in Earth system models. Soil maps show the succession of soil types as a function of soil-forming factors and define occurrence of slow permeable subsurface horizons that may induce seasonal lateral flow patterns, or possible surface soil crusting causing runoff. These considerations are increasingly relevant in the light of appropriate model benchmarking. Remote and proximal sensing and in situ monitoring devices are now widely available, allowing for much better model benchmarking than was possible in the past. Such benchmarking is, of course, essential to further model development and guiding future soil observation collection (Mishra et al., 2017; Mishra & Riley, 2014).

Final challenge for methodological improvement that we address here is the integration of PTFs as knowledge rules for different processes. Parameterization for complex models can imply different PTFs, which

needs precaution as they can refer to same basic properties, but offers also opportunities of integration. Earth system models incorporate and integrate many different processes, and the applied modules/rules/algorithms for the different processes have to be carefully and consistently parameterized, incorporating existing field knowledge and validation (Luo et al., 2016).

3.1. Extrapolation

A PTF user has to select the PTF to use. Many studies were conducted to test different PTFs for predicting soil water retention. However, we are spoiled for choice as a result of the large number of available PTFs (Gijsman et al., 2002). Probably due to this, current LSMs mostly use default soil parameters, which generally do not represent spatial variability (Kishné et al., 2017; Mishra & Riley, 2015). Kishné et al. (2017) found that 95% of the default soil parameters in the model were significantly different from the region-specific observations.

Besides picking a PTF based on its performance, an equally important factor, usually overlooked, is how representative the training data are on the domain of application. The data used to generate a PTF represents soil within a context (e.g., spatial or variables dimension); hence, it has been recommended that a given PTF should not be extrapolated beyond the geomorphological region or soil type from which it was developed (McBratney et al., 2002), as they may lose their validity (Minasny et al., 1999). It has been reported that general PTFs yield poor results when applied in different pedogenetic environments. Casanova et al. (2016) compared laboratory measured bulk density from alluvial soils from Chile with the predictions of 10 published PTFs and highlighted the need to develop local models. Nanko et al. (2014) reported similar results for forest soils in Japan. As a result, many PTFs have been developed for specific regions (Cosby et al., 1984, for USA, and Wösten et al., 1999 and Tóth et al., 2015, for Europe), implicitly defining a spatial context, or more explicitly like the hydraulic parameter maps generated with PTFs by Marthews et al. (2014). Pringle et al. (2007) stressed that when the predictions of PTFs are distributed in space, a spatial evaluation should be performed. They found that the correlation between predicted and observed values varied in space when evaluating four PTFs used to predict soil hydraulic properties.

Nemes (2015) noted that while the rule of thumb is that a PTF should not be used outside the given geographic area, a difference in the geographic area is likely not the true reason for a PTF to fail. It is rather the similarities or differences between the development and application data in data range, as well as the underlying correlation patterns that will determine if a PTF will perform well or fail. However, usually, there is very little general information and metadata published about the respective data sets.

It is possible to measure the suitability of PTFs from different standpoints. McBratney et al. (2002) proposed the use of fuzzy k-means as a measure to describe the training data. Tranter et al. (2010) expanded the idea by using the clustering information to estimate the uncertainty and also assessed whether a sample is within the training data domain, penalizing the prediction uncertainty when the sample is different to the training set. With this approach, the applicability of the PTF is ultimately determined by the uncertainty level of the predictions. The nearest-neighbor-type techniques can also produce a similar functionality if, at the time of calculating the distance metric of a sample from each of the training samples, the distance metrics are intercepted and evaluated. Some aspects of working with various data patterns have been studied in such context by Nemes, Rawls, Pachepsky, and Van Genuchten (2006).

More generally, most of the possible extrapolations that a PTF can generate are related to how well we are able to describe the conditions in which the PTF has been developed. Not only having an idea of the training data is important but also the soil information that is not explicitly used in the model. For example, in certain PTFs to predict water retention, bulk density might not be an important predictor, but information about its magnitude can be an indicator of compaction, which affects porous space and ultimately water retention. Even if complementary measurements are unavailable, information like management attributes can be used as a proxy to understand the pedological context of the training data. We encourage not only the provision of this PTF metadata (including timing and conditions of sampling) but also information about field and laboratory methodology used to obtain the data. The EU-HYDI (Weynants et al., 2013) gives an example of including detailed information on such context of data collection. This idea was also proposed by McBratney et al. (2011), but it is yet to be widely adopted.

PTFs are usually used as inputs into process-based simulation models (Jong & Bootsma, 1996; Mayr & Jarvis, 1999; Young et al., 2002), digital soil mapping routines (Behrens & Scholten, 2006; Marthews et al., 2014; Mishra & Riley, 2012), or even to generate other pedotransfer function (Morris, 2015).

In all applications, the assessment of uncertainty in PTF predictions is crucial. For example, in soil carbon stock assessment, where bulk density is usually not measured, PTFs for bulk density are required and can be the main source of uncertainty (Hollis et al., 2012). PTF uncertainty should be propagated through subsequent models, so that it is quantitatively represented. While many PTFs predictions have been generated, it has not been a general practice to provide uncertainty levels for them. The minimum requirement can be a general measure of error like root-mean-square error. A more advanced approach is to provide uncertainty information per (predicted) point following Tranter et al. (2010). Alternative uncertainty metrics can also be provided by the use of techniques that involve multiple models by design (ensemble modeling) or that can easily implement data resampling (e.g., bootstrapping).

While many PTFs have been developed at different places using different algorithms (Jana et al., 2007; Schaap et al., 2004), less effort has been put into using published PTFs more efficiently. Guber et al. (2009) suggested the use of all available PTFs in a multimodel prediction technique. They used 19 published PTFs as inputs in Richards' (1931) soil water flow equation, the output of the 19 simulations were then combined to obtain a more optimal soil water prediction. The challenge in this type of ensemble method is how to calibrate and use appropriate weighting for each of the PTFs to obtain an optimal prediction.

McBratney et al. (2002) proposed a soil inference system that would match the available input with the most appropriate PTF to predict properties with the lowest uncertainty. The soil inference system was proposed as a way of collecting and making better use of pedotransfer functions that have been abundantly generated. McBratney et al. (2002) demonstrated the first approach toward building a soil inference system (SINFERS). It had two essentially new features; first, it contained a suite of published pedotransfer functions. The output of one PTF can act as the input to other functions (if no measured data are available). Second, the uncertainties in estimates were inputs and the uncertainties of subsequent calculations are performed. The input consists of the essential soil properties. The inference engine will predict all possible soil properties using all available combinations of inputs and PTFs and will select the combination that leads to a prediction with the minimum variance. There have been some attempts at pattern matching of PTFs using a distance metric (Tranter et al., 2009) or nearest-neighbor algorithms (Nemes, Rawls, & Pachepsky, 2006). However, there have been no research applications that do what SINFERS aims to do, to build a system that would chain the PTF predictions together while accounting for uncertainty. The main benefit of the SINFERS approach is that, given a minimum amount of input, it provides the maximum soil property data and expert interpretation of that data, providing soil science expertise as a service. Morris (2015) built an expert system software, which uses rules to select appropriate PTFs and predicts new property values and error estimates. SINFERS can use the estimated property values as new inputs, which can trigger more matching patterns and more PTFs to "fire" cyclically until the knowledge base is exhausted and SINFERS has inferred everything it can about what it was originally given. The next logical step after accounting for cumulated uncertainty will be to test how those translate in a mapping or numerical simulation application.

In support to the general applicability, PTFs have been reported to be rather flexible with respect to local calibration, especially for functional behavior of a system, illustrated for soil hydraulic characterization calibrated with information from digital elevation models (Romano & Palladino, 2002; Romano, 2004). Digital soil mapping for modeling the spatial and temporal variability in soil properties can further be achieved with pedotransfer functions applying auxiliary data from landscape and terrain analysis, remote or proximal sensing, geostatistics, etc. (Mulla, 2012). The use of spatially referenced soil profile description data and environmental variables (topography, climate, and land cover) through different regression approaches shows possibilities for improvement of PTFs (Mohanty, 2013) to predict the spatial variability of soil properties, as has been explored for carbon stocks by Mishra and Riley (2012) and for determining active layer thickness of soils in permafrost regions (Mishra & Riley, 2014).

In spite of the aforementioned efforts, still, there is a huge knowledge gap, especially for specific, often under-represented soil systems, such as saline (Tóth et al., 2012) and calcareous soils (Khodaverdiloo et al., 2011), volcanic ash soils (Nanko et al., 2014), peat soils (Hallema et al., 2015; Rudiyanto et al., 2016), paddy soils, soils with well-expressed shrink-swell behavior (Patil et al., 2012), and soils affected by freeze-thaw cycles

(Yi et al., 2013). Difficulties with determining hydraulic and geochemical behavior for deriving PTFs in these soil types are due to processes such as temporary nonequilibrium conditions and uncertain estimates of the actual field conditions due to measurement uncertainty (e.g., Dettmann & Bechtold, 2016).

3.2. Scaling

The spatial heterogeneity of land surfaces affects energy, moisture, and greenhouse gas exchanges with the atmosphere. However, representing the heterogeneity of terrestrial hydrological and biogeochemical processes in ESMs remains a critical scientific challenge (Mishra & Riley, 2015). Most PTFs have been calibrated from point source data and assume absence of spatial correlation. In digital soil mapping the main interest is in estimating the spatial distribution of soil properties. Pringle et al. (2007) recommended that an investigator who wishes to apply a PTF in a spatially distributed manner first has to establish the spatial scales relevant to their particular study site. Following this, the investigator must ascertain whether these spatial scales correspond to those that are adequately predicted by the available PTFs. They proposed three aspects of performance in the evaluation of a spatially distributed PTF: (i) the correlation of observed and predicted quantities across different spatial scales, (ii) the reproduction of observed variance across different spatial scales, and (iii) the spatial pattern of the model error. For an example of predicting water retention across a 5 km transect, Pringle et al. (2007) showed that the tested PTFs performed quite well in reproducing a general spatial pattern of soil water retention. However, the same study reported that the magnitude of observed variance was well underestimated, exemplifying how PTF estimates tend to be smoother than the original data.

A formidable challenge is PTF development for coarse-scale soil modeling, such as for LSMs. Soil parameters in these models cannot be measured, and the efficiency of PTFs can be evaluated only in terms of their utility (Gutmann & Small, 2007; Shen et al., 2014). The general problem of change in spatial resolution of the input data by aggregating small-scale information and the resulting output uncertainty for various model states was reported, for example, by Cale et al. (1983), Rastetter et al. (1992), Hoffmann et al. (2016), and Kuhnert et al. (2016). Besides presenting spatially dependent variability, PTFs may also have scale dependence (Pringle et al., 2007). This characteristic has been widely reported in hydropedology where upscaling PTFs is needed to model watershed processes (Pachepsky & Rawls, 2003; Pachepsky et al., 2006). A potential solution to get an adequate representation of coarse-scale parameters might be the fitting of the target function through all subgrid function representations. Montzka et al. (2017a) generated a global map of soil hydraulic parameters for a coarse 0.25° grid based on the fine SoilGrids 1 km database by fitting a single water retention curve through all subpixel retention curves. Dai et al. (2013) have prepared a soil hydraulic map of China of seven soil depths of up to 1.38 m at 1 km resolution using several widely accepted point and parametric PTFs applied on the soil map of China. Tóth et al. (2017) have calculated soil water retention, hydraulic conductivity at certain matric potential values, and MvG parameters at seven soil depths of up to 2 m depth at 250 m resolution for Europe with the EU-HYDI PTFs of Tóth et al. (2015) based on SoilGrids 250 m (Hengl et al., 2017). Chaney et al. (2016) estimated van Genuchten hydraulic parameters at various depths (from 0 to 2 m) based on PTFs input of sand, clay, bulk density, and water content at −33 and −1500 kPa. The resulting soil map of the contiguous USA at a resolution of 30 m, with data available at www.polaris.earth, is probably the highest-resolution soil hydraulic map at (sub)continental domain.

Using these approaches, PTFs can be applied on high-resolution data sets of basic soil properties. It needs to be stressed that SoilGrids and other global products were derived from point measurement. While the grid spacing is at either 1 km or 250 m, its spatial support is still at a point. Thus PTFs (derived from point observations) are mainly applied to rasters of regional, continental, or global maps, which is not an upscaling issue.

Also, the coarse spatial scale often assumes a coarse temporal support, which requires an understanding of how to include other environmental variables in PTFs, such as weather and management attributes. Such incorporation might be derived with nearest-neighbor methods, with sets of global parameters obtained from the k-nearest-neighboring region using various metrics to take into account the uncertainty of its parameterization (Samaniego, Kumar, & Attinger, 2010). Other proposed approaches were recently developed to tackle the deficiencies in existing models such as overparameterization, the lack of an effective technique to integrate the spatial heterogeneity of physiographic characteristics, and the nontransferability of parameters

across scales and locations. A multiscale parameter regionalization technique is proposed as a way to address these issues simultaneously (Samaniego, Bárdossy, & Kumar, 2010). In this multiscale parameter regionalization (Samaniego, Bárdossy, & Kumar, 2010) model parameters instead of basin predictors are aggregated first. Other proposed ameliorations lie in the use of geographically weighted regression enabling to enter scale-specific relationships both in the development and application of PTFs.

3.3. Integration

There is an urgent need to determine combinations of PTFs, together with upscaling procedures that can lead to the derivation of suitable coarse-scale soil model parameters. However, simple regressions and any statistically derived inference may provide wrong conclusions when combined. Several recent papers report on combinations of PTFs; Tang et al. (2015) demonstrated the coupled modeling of root and soil water transport being more robust than the sequentially coupled modeling. These authors also quantified the uncertainty related to the use of different PTFs in global evapotranspiration (ET) estimation. The strength of coupled estimations was also shown by Huang et al. (2016) who present a coupled carbon-nitrogen Earth ecosystem model with robust coupled estimations at global scale for total gross ecosystem productivity, ecosystem respiration (R_e), net ecosystem productivity, net primary productivity, latent heat, sensible heat (H), soil organic carbon (SOC), and total vegetation biomass. Same way, biogeochemical models of competition for soil N that estimate multiple processes simultaneously match the observed patterns of N losses better than models based on sequential competition (Niu et al., 2016).

Examples of exploration of relationships for process parameterization are metaanalyses. For example, Fóti et al. (2016) defined the correlations between environmental gradients and spatial patterns of soil respiration R_s, soil water content, and soil temperature. Other possible ways to develop improved empirical relationships are by data mining techniques as explored by Jarvis et al. (2013) in collating a global database of hydraulic conductivity measured by tension infiltrometers under field conditions. Their resulting PTFs contrast markedly with the existing to estimate K_r. For example, saturated hydraulic conductivity, K_s, in the topsoil (<0.3 m depth) was found to be weakly related to texture. Instead, they inferred stronger relationships between K_s and bulk density, organic carbon content, and land use, which they determined more rigorously afterward with boosted regression trees (Jorda et al., 2015). With spatial data mining the scaling problems for application of estimations for soil, ecosystem, and climate can also be solved (Feoli et al., 2017).

Another recent technique that has merits in this respect is ensemble modeling—that is, the use of a number of models in combination. This technique is a natural part of weather and climate modeling today, yet it is less used in the prediction of soil properties (Baker & Ellison, 2008b). Ensemble modeling carries a number of benefits and potential over the use of a single model. Models can differ in their theory and structure and also in the information that they require. As a result, their sensitivity and scale of support may also differ. The use of ensemble modeling is easy to justify if it is difficult to determine which, if any, single model may be superior to others. In ensemble modeling, the main aim is not to make the single model perfect but to capture the trend that multiple models agree on. The ensemble will amplify trends that are common among models, while by-chance predictions will be softened. The outputs, therefore, can be interpreted—qualitatively or quantitatively—as a measure of uncertainty. In the context of integrated Earth system models, the represented complex processes—integrating physical and biochemical processes typically—can be covered by a number of models with strongly varying concept and structure. Here lies an opportunity to construct ensemble models entering different PTF-based parameterizations.

Up to now, in soils related predictions two different types of ensemble models have been explored. Guber et al. (2009) used 19 published PTFs in an ensemble prediction scheme to parameterize a Richards' soil water flow model. In their scheme, they used different models that required different sources and levels of input and that had different structure. A more popular approach—and one that is more straightforward to implement—is the use of several schemes to resample data of the main data pool and use those to develop a given number of predictive models of the same structure, which are then statistically pooled to give a prediction—optionally with a measure of uncertainty. Such schemes include bagging and bootstrapping and have been used by some authors in this field (e.g., Baker & Ellison, 2008b; Nemes et al., 2010; Schaap et al., 2001).

Table 4
Selection of Widely Applied PTFs for the Parameterization of Moisture Retention (MRC) and Hydraulic Conductivity (HCC) Curves of the Mualem-van Genuchten (MvG) and Brooks-Corey (BC) Models

PTF	Source	Region	Model	No. of samples	Sand	Silt	Input clay	Bulk density	Organic matter	Other
1	Baumer (1992)	USA		18,000	+	−	+	+	+	
2	Pachepsky et al. (1982)	Hungary	BC	230	+	+	+	+	−	
3	Clapp and Hornberger (1978)	USA	BC	1,446	(+)	(+)	(+)	−		Soil classes
4	Cosby et al. (1984)	USA	BC	1,448	+	+	+	−	−	
5	Rawls et al. (1982)	USA	BC	5,320	(+)	(+)	(+)	−	(+)	Soil classes
6	Rawls and Brakensiek (1985)	USA	BC	5,320	+	−	+	+	−	
7	Saxton et al. (1986)	USA	BC	5,320	+	−	+	−	−	
8	Carsel and Parrish (1988)	USA	MvG	5,097–5,693	(+)	(+)	(+)	−	−	
9	Vereecken et al. (1989)	Belgium	MvG	182	+	−	+	+	+	
10	Williams et al. (1992)	Australia	BC	196	+	−	+	−	+	
11	Wösten et al. (1999)	Europe	MvG	1,136–2,894	+	+	+	+	+	T/S
12	Schaap et al. (2004) (ROSETTA)	USA and Europe	MvG	2,134/1,306	+	+	+	+	−	
13	Weynants et al. (2015), Weihermüller et al. (2017)	Belgium	MvG	136 (HCC), 166 (MRC)	+	−	+	+	+	
14	Tóth et al. (2015)	Europe	MvG	134–6,074	+	+	+	+	+	T/S, pH, CEC, and CaCO3
15	Zhang and Schaap (2017)	USA and Europe	MvG	2,134/1,306	+	+	+	+	−	

4. PTFs in Earth System Sciences

The process formulations of models in Earth system sciences, and more specifically of ESMs with regard to their Land Surface Model (LSM) compartments, rely on equations that capture as much as possible the underlying processes describing the complex nature of the Earth system. Hereby, the different compartments of the LSM require equations for, for example, radiation exchange between the land surface and atmosphere, root water uptake by vegetation, and mathematical descriptions for hydraulic as well as thermal flow processes in the soil compartment. Poor descriptions of soil physical processes and/or related parameter choice have been listed as one of the possible explanations for the lack of consensus on the correct magnitude of the land surface-atmosphere coupling strength in general circulation models. This coupling strength plays a role in closing the hydrological cycle at various time scales, which the model modulates via memories in the terrestrial water balance (see Koster et al., 2000; Seneviratne et al., 2006). A number of studies have illustrated the considerable sensitivity of the hydrological cycle (and related energy fluxes) to soil hydraulic properties on the global scale (e.g., Ducharne et al., 2000; Oleson et al., 2013), regional scale, for example, on the skill of the UK numerical weather prediction model ("Unified Model," see Dharssi et al., 2009) and field scale (Brimelow et al., 2010). Most LSMs mainly use default soil parameters, oversimplifying their influence. Kishné et al. (2017) found that using measured and PTF-estimated water content at field capacity and wilting point yielded a 35 to 76% decrease in plant available water compared to default model settings.

For the parameterization of hydraulic and thermal properties of the soil, LSMs greatly rely on PTFs because the models they are part of are generally run on regional to global scale, where measured data for the water retention and hydraulic conductivity as well as thermal property functions are not available. For the estimation of the shape of the hydraulic and thermal functions the different LSMs mainly use embedded PTFs, which are fed by external data from regional to global soil maps. These inputs can be discrete textural information (percentage of sand, silt, and clay content), organic carbon content, and bulk density (among others) or soil class information (e.g., sand, loam, and sandy clay).

This part of the review is dedicated to summarizing the recent developments of PTFs employed for calculation of water, solute, and heat flow, as used in LSMs. PTFs for the hydraulic properties (section 4.1) are mainly needed in the LSM to help calculate the water movement in the vadose zone, for example, with Richards' equation (see equation (17) and Table 5) and related soil moisture profiles. Section 4.2 focuses on the use of PTFs to describe solute transport processes in soils. Soil thermal properties (section 4.3) used in LSMs have a large influence on soil heat flux, and hence on the available energy at the land surface (in particular for sparsely vegetated surfaces) and also on the soil water balance via soil moisture-ice phase changes, via latent heat release and temperature effects (see, e.g., Peters-Lidard et al., 1998, one of the few studies available on this topic).

4.1. PTFs of Water Flow
4.1.1. Soil Water Content and Flows
A major field of application of PTFs in Earth system sciences is the prediction of soil hydraulic parameters (see Table 4) that are key in the prediction of the water and energy balance of the land surface-vadose zone system and control, for example, the transport of solutes in the subsurface environment. The parameter estimation approach introduced in section 2.2 is used in most of vadose zone models based on Richards' equation (equation (17)) to describe water flow. Early PTFs were mainly regression equations that either estimate soil water content at fixed pressure head (e.g., field capacity or permanent wilting point) from texture, bulk density, and organic matter (Gupta & Larson, 1979) or parameter estimation methods that relate soil properties to the parameters in simple power law equations such as the Brooks-Corey or Campbell equation used to describe

Table 5
Models for the Description of Moisture Retention (MRC) and Hydraulic Conductivity (HCC) Curve (After Assouline & Or, 2013; Kosugi et al., 2002; Kutílek and Nielsen, 1994; Lal & Shukla, 2004; Li et al., 2014; Vereecken et al., 2010)

Model name	Expression		Parameters
	Moisture retention curve		
Burdine (1953)	$S = \frac{1}{(1+(\alpha h)^n)^{1-2/n}}$		S: degree of saturation $S = \frac{\theta-\theta_r}{\theta_s-\theta_r}$ h: matric potential; α, n: fitting parameters
Brooks and Corey (1964)	$S = \left(\frac{h_A}{h}\right)\lambda S = 1$	if $h < h_A$ if $h \geq h_A$	h_A: air entry matric potential; λ: fitting parameter related to tortuosity and connectivity of the pores
Brutsaert (1966)	$S = \frac{1}{(1+(h\,\alpha)^n)}$		
Campbell (1974)	$\frac{\theta}{\theta_s} = \left(\frac{h_A}{h}\right)\lambda\frac{\theta}{\theta_s} = 1$	if $h < h_A$ if $h \geq h_A$	
Mualem (1976)	$S(1 + (\alpha h)^n)^{-1 + 1/n}$		
van Genuchten (1980)	$S = [1 + (\alpha h)^n]^{-m}$		m : fitting parameter $m = 1 - 1\backslash n$
Fredlund and Xing (1994)	$S = (\ln(e + (h/\alpha)^n))^{-m}$		α, n, m: constant; e: base of natural logarithm
Kosugi (1996)	$S = \frac{1}{2}(\theta_r - \theta_s)\,\text{erfc}\left[\frac{\ln\left(\frac{h}{h_{0.5}}\right)}{\sigma\sqrt{2}}\right]$		θ_r: residual water content; θ_s: saturated water content; σ: constant, erfc: complementary error function, $h_{0.5}$: median matric potential, corresponding to the median pore radius
	Hydraulic conductivity curve		
Gardner (1958)	$K(\theta) = K_S \exp\left[c(h - h_A)\right]$		θ: water content; a, b, m are empirical coefficients
Brooks and Corey (1964)	$K(\theta) = K_S \left(\frac{h}{h_b}\right)^{-n} K(\theta) = K_S$	if $h < h_A$ if $h \geq h_A$	K_S: hydraulic conductivity at saturation
Mualem and Dagan (1978)	$K(\theta) = K_S S^l \left(\frac{\int_0^\theta 1\ h^{1+b}\mathrm{d}\theta}{\int_0^\theta \text{sat}\ 1\ h^{1+b}\mathrm{d}\theta}\right)^2$		b: fitting parameter related to tortuosity; l: fitting parameter related to pore size

the moisture retention characteristic and the unsaturated hydraulic conductivity (Clapp & Hornberger, 1978; Saxton et al., 1986; Saxton & Rawls, 2006).

For the interpolation of measured water retention and hydraulic conductivity data for the whole tension or suction range, mathematical expressions describing the full moisture retention (MRC) and hydraulic conductivity curve (HCC) are used (Smettem et al., 2004). Table 5 lists the most widely used functions for the description of the MRC and HCC. For consistency in numerical soil hydraulic modeling a coupling of the MRC with the HCC is needed, for which most frequently the model of Mualem-van Genuchten (MvG) is applied (equations (13)–(16)).

$$\theta(h) = \theta_r + \frac{\theta_s - \theta_r}{[1 + (\alpha h)^n]^m} \tag{13}$$

$$m = 1 - 1/n \tag{14}$$

$$K(S_e) = K_0 S_e^L \left[1 - \left(1 - S_e^{1/m}\right)^m\right]^2, \tag{15}$$

$$S_e(h) = \frac{\theta(h) - \theta_r}{\theta_s - \theta_r}, \tag{16}$$

where $\theta(h)$ is the water content of the soil (cm^3 cm^{-3}) at a given matric potential value (centimeter of water column); θ_r is the residual water content (cm^3 cm^{-3}); θ_s is the saturated water content (cm^3 cm^{-3}); and α (cm^{-1}), n (−), and m (−) are fitting parameters; $K(S_e)$ is the soil hydraulic conductivity (cm d^{-1}) at certain effective saturation (−); K_0 is the hydraulic conductivity acting as a matching point at saturation (cm d^{-1}) and L is a shape parameter related to tortuosity of the pore space (−).

The models were further improved to decrease the uncertainty close to saturation, for example, by Schaap and van Genuchten (2006), Ippisch et al. (2006), and Jarvis (2008). MRC and HCC models have been recently reviewed by Assouline and Or (2013). Weynants et al. (2009) found that MRC model of Ippisch et al. (2006) coupled with HCC of Mualem (1976) did not improve the description of water retention but significantly improved the simulation of hydraulic conductivity near saturation. Improvement of

the MvG model can be achieved at low matric potential values as also mentioned by Tuller and Or (2001) who highlighted the importance of film flow contribution in unsaturated hydraulic conductivity. Soil water flow is a process of transient nature. Using pressure-based soil water contents to determine characteristic points of the water retention curve like field capacity implies the risk to neglect this transient behavior. Several authors discussed this problem critically and developed methods to describe the field capacity using dynamic approaches (e.g., Assouline & Or, 2014; Nachabe, 1998; Meyer & Gee, 1999; Twarakavi et al., 2009; Zacharias & Bohne, 2008).

The Brooks and Corey (1964) and Campbell (1974) models are still in use, but the Mualem-van Genuchten (MvG) model is now one of the most used models to describe the MRC and HCC in vadose zone research as it has the largest flexibility in describing a large range of moisture retention curve shapes and is currently implemented in several vadose zone models (e.g., HYDRUS and SWAP) and LSMs (e.g., ORCHIDEE, see section 4.1.3). Yet this model, as all other models developed to estimate water retention and unsaturated hydraulic conductivity characteristics, relies heavily on the concept of flow in capillary pores and rigid porous media. Recent developments in predicting the parameters of the MvG model using pedotransfer functions have been reviewed by Vereecken et al. (2010). In addition, several reviews analyze the state of the art in developing PTFs (Pachepsky & Rawls, 2004; Wösten et al., 2001). Various databases are now available that provide soil hydraulic information as well as basic soil properties to derive PTFs, such as HYPRES (Wösten et al., 1999), UNSODA (Leij et al., 1996; Nemes et al., 2001), ROSETTA (Schaap et al., 2001), IGBP-DIS soil data set (Tempel et al., 1996), EU-HYDI (Weynants et al., 2013), and Ahuja database (Schaap & Leij, 1998a). Nemes (2011) provided an overview of the pros and cons and availability of most independent databases of international interest—available at the time—that contain significant quality and quantity of soil physical and hydraulic information of point samples. Based on a literature analysis of published data, Vereecken et al. (2010) showed that the root-mean-square error (RMSE) of the fitted MvG model to moisture retention characteristic data is of the order of 0.017 ($cm^3\ cm^{-3}$). The accuracy of most PTFs that are used to generate moisture retention characteristics is now approaching this value but still with a scope for improvements as RMSEs are still 2 to 5 times larger than RMSEs values obtained from fitting the MvG model to measured data. The lowest RMSE values were obtained when organic matter (or organic carbon) was included in the equation in addition to texture and bulk density (Cornelis et al., 2001; Tietje & Tapkenhinrichs, 1993). There have been attempts to omit the use of organic matter as predictor in the PTFs (e.g, Campbell & Shiozawa, 1992; Nemes et al., 2003; Oosterveld & Chang, 1980; Zacharias & Wessolek, 2007) because in general bulk density is correlated with soil organic matter content. On the other hand, bulk density can be a very dynamic soil property influenced by other factors than organic carbon content, such as agricultural management practices, weathering, drainage, or biological activities. Forgoing the organic matter content in the list of PTF predictors can help to achieve a better representation of this dynamic character in the prediction of soil hydraulic properties (Zacharias & Wessolek, 2007).

The most common PTFs for predicting hydraulic conductivity either use soil properties as predictors using regression statistics (e.g., Cosby et al., 1984; Saxton et al., 1986; Vereecken et al., 1990; Wösten et al., 1999) or apply physical-empirical relationships relating particle size distribution and hydraulic conductivity using the concept of effective porosity and effective pore radii (e.g., Ahuja et al., 1984; Campbell, 1974; Millington & Quirk, 1959; Timlin et al., 1999).

For the hydraulic conductivity curve the RMSE ratio of the estimated hydraulic conductivity from PTFs to the direct fit was about two. The most accurate PTFs to estimate the complete hydraulic conductivity curve described by MvG used the estimated saturated hydraulic conductivity rather than the measured one and an estimated tortuosity parameter, L, rather than the standard value of 0.5 proposed by Mualem (1976). The reason for the preference for estimated values of hydraulic conductivity over measured ones in general is the inherent scale dependency of this parameter (Gelhar, 1986; Roth, 2008). The value for hydraulic conductivity is dependent upon the pore geometry at the scale of interest (Nieman & Rovey, 2009; Sánchez-Vila et al., 1996; Schulze-Makuch et al., 1999) and may show atmospheric dependence (Oosterwoud et al., 2017) seasonal effects (Suwardji & Eberbach, 1998; Farkas et al., 2006; Borman & Klaasen, 2008), and its quantification methods can produce as much variation as the other mentioned factors (Fodor et al., 2011). An approach similar to that used in groundwater hydrology, where the heterogeneity in the hydraulic conductivity is remediated with defining an effective conductivity,

may be useful for use of hydraulic conductivity PTFs in ESMs. Bevington et al. (2016) have applied factorial kriging analysis for identifying the processes that affect soil variables by removing their scale dependency.

Vereecken et al. (2010) identified a series of challenges in the development of hydraulic PTFs. These included the following: (1) improving measurements of soil hydraulic properties used to derive the PTFs, (2) the development of comprehensive databases and new data exploration techniques, (3) the definition of comprehensive metadata, (4) adding relevant predictors such as structural properties, and (5) the development of time-dependent PTFs. Alternative approaches to the concept of capillary flow as a basis for deriving models of unsaturated hydraulic conductivity have been proposed by, for example, Tuller and Or (2001). The approach of Tuller and Or (2001) allows accounting for film and corner flow, important features under dry conditions. The same idea has been also taken up by Peters and Durner (2008). In addition, we need to move beyond unimodal soil hydraulic functions such as MvG if we want to include the impact of soil structure on these functions and water flow. Fitting multimodal models (e.g., multiporosity) for soil hydraulic properties is a prerequisite for incorporating soil structural information in PTFs. However, for multimodal models no PTFs have been developed yet. Most PTFs that are nowadays being used in vadose zone research, as well as in weather and climate research and prediction, strongly rely on the use of texture as predictor. Yet the inclusion of bulk density, organic matter, and soil chemical properties such as pH, calcium carbonate content, or cation exchange capacity may greatly improve the prediction capabilities of PTFs (Tóth et al., 2015; Vereecken et al., 2010). Recently, a global soil data set has been published that provides information on up to 250 m resolution not only on texture but also basic further soil physical and chemical properties along with the depth of soil profiles based on the GlobalSoilMap specifications (Arrouays et al., 2014; SoilGrids 250 m; Hengl et al., 2017). These data may provide an excellent basis for the derivation of global depth-dependent soil hydraulic parameters as well as their statistical properties (e.g., variance and spatial correlation). This may, for instance, prove to be very valuable in relating empirical infiltration parameters to basic soil properties or derive information that is key to quantify infiltration (e.g., depth to bedrock). Quantification of infiltration in LSMs is either based on empirical equations or analytical numerical solutions of water flow models based on either Richards or bucket-type equation. The parameters needed for this purpose include the hydraulic conductivity and the soil moisture characteristic. Therefore, existing PTFs can be used to predict and quantify the infiltration process including its spatial variability. Improvements can also be made in the way the infiltration process is described as typically the role of structure is neglected. These improvements can be obtained by developing PTFs of soil hydraulic properties that account for the role of macropores, stoniness, and by considering concepts like dual permeability models or the multimodal models.

4.1.2. Root Zone Hydraulic Processes

Plant roots have an important impact on key soil processes such as water flow, solute transport, carbon fluxes, and biogeochemical processes. Soil process and land surface models are therefore very sensitive to the parameterization of the root zone (Hinsinger et al., 2011; Javaux et al., 2013). The first crucial parameterization concerns root water uptake—and the resulting transpiration flux. Typically, root water uptake is accounted for in Richards' equation with a sink-source term S ($L^3 L^{-3} T^{-1}$):

$$\frac{\partial \theta}{\partial t} = \nabla \cdot [K(h) \nabla h] + \frac{\partial K(h)}{\partial z} + S, \tag{17}$$

where z denotes the vertical coordinate (L), h the soil water pressure head (L), and $K(h)$ the soil hydraulic conductivity tensor (L T^{-1}). The two first terms describe the water flow redistribution between layers or soil locations, while the third one describes the water uptake by plant roots ($S < 0$) or root exudation ($S > 0$). This sink term S is mostly estimated based on root length density distribution, and the soil hydraulic resistance distribution often represented by a PTF of the bulk soil matric and/or osmotic potential. For root water uptake and the nonequilibrium water retention in the root zone, specific PTFs have been developed based on soil matric potential and dry weight of soil (Schwartz et al., 2016).

The second crucial parameterization concerns root nutrient uptake for which models describe uptake processes in relation to soil hydraulic properties, as by Michaelis-Menten kinetics (Darrah et al., 2006). For nutrients of low mobility such as phosphate, specific parameterization exists for uptake models. In addition, these nutrient uptake models have been coupled with soil water flow models (Somma et al., 1998; Roose & Fowler, 2004). An additional specific rhizosphere related question is the impact of soil physical

properties on roots and vice versa. How macropores are used by roots and how roots create macropores or induce compaction are still challenging questions, which only start to be included in models (Kautz, 2014). However, until now these improved descriptions have not yet been sufficiently incorporated into larger-scale models (Hinsinger et al., 2011; Vereecken et al., 2016). Recent initiatives in this context already include soil resistance, plant root distribution, and climatic demand, to scale up to the macroscale (Javaux et al., 2013).

Root zone depth and root density are the structural parameters that have to be estimated, which are intimately linked to basic soil properties such as soil texture, structure, and organic matter content. Subsequently, the activity of the roots within the root zone (i.e., water and nutrient uptake, respiration, and root exudation) is defined to be proportional to root density. However, root density can be defined in terms of several parameters: root length density, root surface density, or root mass density. Furthermore, root activity can be attributed to different types of roots, which are assigned to different functions: for example, fine roots have a water/nutrient uptake function and thick roots have storage and transport functions. The functional classification of roots is still a matter of intensive research (Rewald et al., 2011). Root physiologists determined root properties and parameters such as root hydraulic conductance and nutrient uptake rates and related these to key structural features of cells and tissues, which in turn can be related to plant functional type classifications. The function of a root segment does depend not only on the properties of that segment but also on the location of the segment in the larger root architecture. Root architecture models describe the structure of plant roots using plant species specific parameters (Dunbabin et al., 2013; Leitner et al., 2010; Pages et al., 2004). When linked to growth models, the dynamics of the root architecture can be represented and coupled to the environmental soil conditions and soil properties. Root growth rate is, for instance, described in these models as a function of soil penetration resistance. "Pedotransfer functions" that relate root growth rate to proxies of penetration resistance, that is, soil bulk density and soil water content, have been derived (Bengough et al., 2006; Gao et al., 2014).

In fully integrated soil-plant models, which couple the description of processes within the plant to soil processes, the function of different roots with different properties and of roots in soil layers with different environmental conditions emerges from the coupled process descriptions. This modeling approach allows, for example, to infer the role of deep roots for water uptake directly (Javaux et al., 2013). However, the application of these fully coupled root-soil models is computationally expensive since processes must be described at the scale of single plant roots. Therefore, these processes are generally parameterized in larger-scale simulation models, with root zone depth and root distribution being the most common input parameters that have to be defined.

To this end, LSMs often define root parameters as a function of the plant functional type or ecosystem. This approach represents the interaction between climate, soil, and vegetation indirectly by the spatial correlation that exists between vegetation types and climate. However, large variations of rooting depth within ecosystems or plant functional types have been observed. Therefore, models have been developed to link root depths with climate and soil type. The first type of model is a statistical model (generalized linear regression models) that predicts the median root depth and 95th percentile of the root depth distribution based on climate and soil data (Schenk & Jackson, 2002, 2005). Climate variables (potential ET, precipitation, and length of the growing season) were found to explain the largest part of the variance in a global data set of rooting depths. Soil texture and the thickness of an organic top layer explained part of the variance of the rooting depth. A close link between root zone water storage, which depends both on the rooting depth and the water storage capacity of the soil, and climate variables could also be inferred from catchment discharge data (Gao et al., 2014). The study of Gao et al. (2014) supported the hypothesis that plant roots develop in such a way that they optimize water and nutrient uptake and survival after droughts at minimal carbon cost. This hypothesis is the basis of models that use optimality criteria to predict root distributions (Guswa, 2008; Kleidon & Heimann, 1998; Schymanski et al., 2008; van Wijk & Bouten, 2001). The key soil parameter in these models is the total plant available water content (estimated from the difference between the water content at field capacity and the wilting point), which is derived from soil texture using pedotransfer functions. Schenk (2008) used a simple soil water balance model to predict the "shallowest model for root water extraction" and found an amazingly good agreement between measured and predicted root distributions.
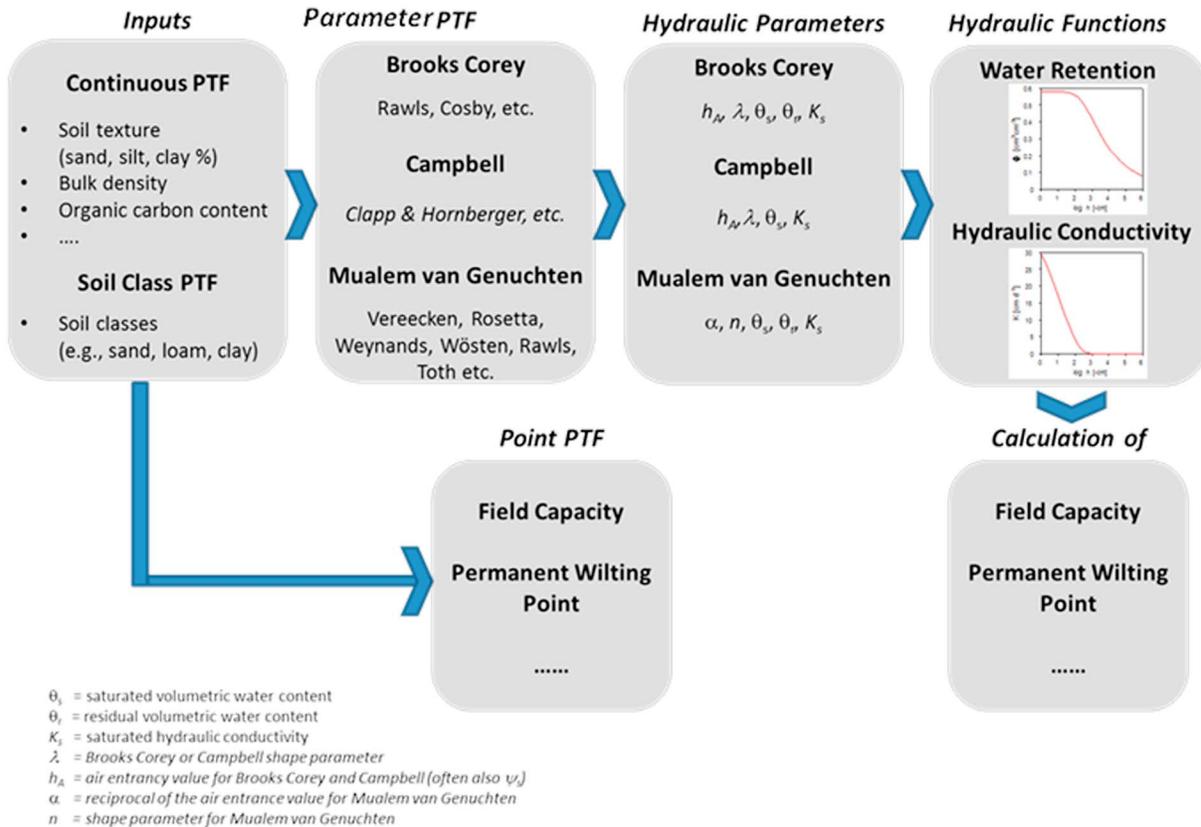
Figure 4. Flow chart for the estimation of hydraulic parameters and the hydraulic functions using pedotransfer functions (PTFs).

Yang et al. (2016) used the model of Guswa to generate global soil root depth maps and found that the prediction of actual evapotranspiration could be improved when using these maps.

4.1.3. Hydraulic Parameterization in LSMs

The different LSMs use different hydraulic PTFs (see Table 4), for the Brooks and Corey (1964), Campbell (1974), or Mualem van Genuchten (1980) models of the soil moisture retention and hydraulic conductivity function. As can be seen from Figure 4 the estimation always starts with input data, which are either continuous soil information (e.g., sand, silt, clay and organic carbon content, and bulk density) or soil class information (e.g., 12 USDA textural classes) or combination of continuous and class/categorical data. In a next step, this information will be fed into the embedded PTF. The output of the PTF will be hydraulic parameters for the corresponding function implemented into the LSM configuration, which are then used to solve the hydraulic functions over the entire range of water contents or pressure heads. This is necessary because estimates of soil moisture content, matric potential, and hydraulic conductivity are required at each time step.

Figure 5a shows a global map of saturated hydraulic conductivity ($\log 10(K_s)$ (cm d^{-1})) as an example for a PTF result applied to global soil data. Here the SoilGrids 1 km data set (Hengl et al., 2014) provides soil texture and bulk density at global 1 km scale based on several soil profile databases and automated mapping of covariates data such as topography, climate, and vegetation information. MvG parameters were estimated by pixelwise application of the ROSETTA PTF (Schaap et al., 2004). Low K_s values can be found in India, the Sahel, the Mediterranean, central Asia, Middle East, the U.S. prairie regions, California, and South Central Canada. Highest K_s are located in Sahara and Arabian Peninsulas, where sandy soils dominate, and also in the upper Amazon basin and in cold climates (Montzka et al., 2017b). WRC and HCC can be estimated from the hydraulic parameter set for each pixel to provide field capacity, wilting point, and soil moisture for the LSM. However, we have to keep in mind that the PTF for estimating K_s was developed from 1,306 soil samples mainly from Europe and North America.
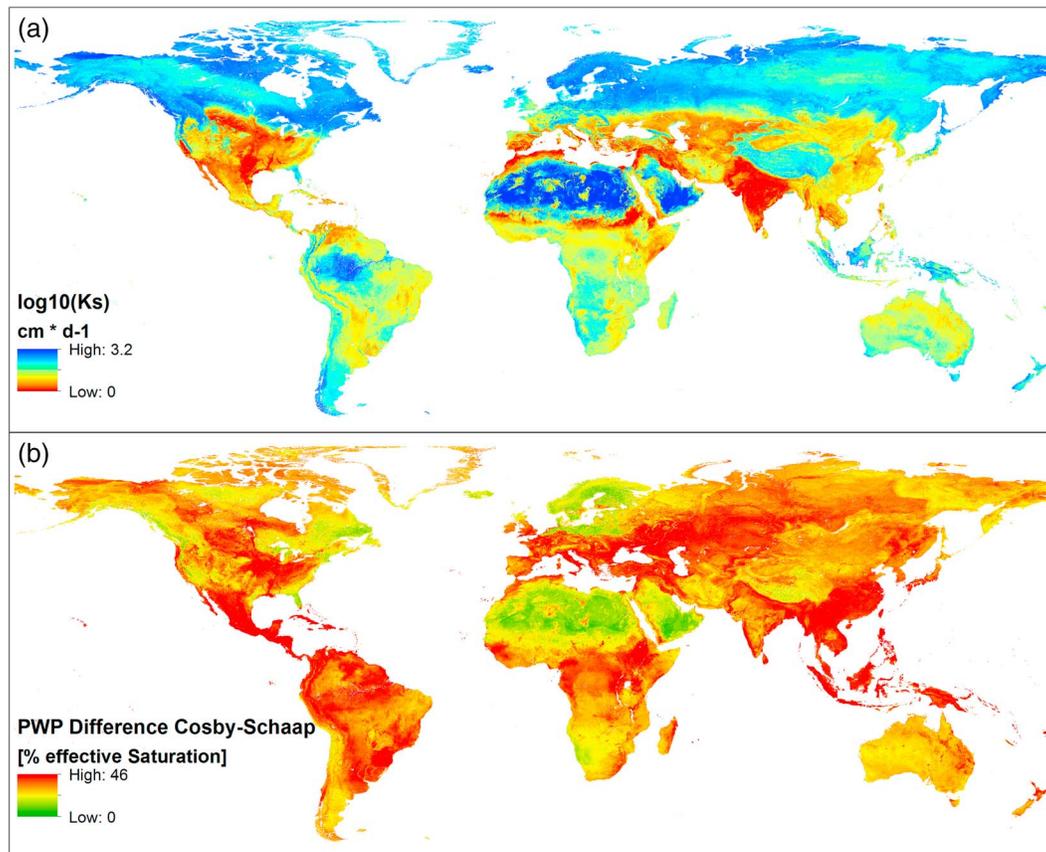
Figure 5. (a) Global map of saturated hydraulic conductivity ($\log_{10}(K_s)$) and (b) differences in permanent wilting point ($h = -15,000$ cm) between the Cosby et al. (1984) PTF with the Brooks-Corey model and the Rosetta PTF with the MvG model in terms of effective saturation (%). Calculations are based on the SoilGrids 1 km data set (Hengl et al., 2014).

For some LSMs, additional point information from the hydraulic functions are needed such as field capacity or permanent wilting point to parameterize, for example, plant water stress functions and carbon turnover rates. The most straightforward way would be to calculate these points from the estimated water retention function. However, some land surface models (for example, ISBA-SURFEX, MPI-HM, and JSBach; see Table 6) calculate these key soil moisture points using a separate embedded PTF (indicated by the point PTF box in Figure 4). The reasons for this "two-way" PTF solution can be found in the development of LSMs from a bucket-type model to a multilayer approach, where originally single retention points such as field capacity and wilting point information were used to describe water flow and other processes. By substituting the bucket model by Richards' equation (equation (17), Richards, 1931) these points were not any longer necessary to predict water flow but still needed for description of other processes related to plant water availability, for example. Instead of solving the water retention curve for predefined points (e.g., wilting point and field capacity) another set of PTFs was introduced. This could potentially cause inconsistencies between soil water flow and plant water uptake, for example. In some cases (e.g., ISBA-SURFEX, MPI-HM, and JSBach; see Table 6) the reason for this approach is related to error reduction with regard to spatial aggregation of the parameters (see Noilhan & Lacarère, 1991).

In order to illustrate potential issues caused by combining different PTFs and different retention models, Figure 5b shows a map of relative differences in permanent wilting point (PWP, $h = -15,000$ cm) between the Cosby PTF (Cosby et al., 1984) with the Brooks-Corey model and the Rosetta PTF (Schaap et al., 2001) with the van Genuchten model. It is obvious that sandy regions such as the deserts and northern Europe show no difference in PWP, whereas in clay-rich soils the differences can reach up to 40% of the effective saturation. The use of different PTFs within one simulation model, even when using this information for different processes, can introduce inconsistencies within the model results and weaken interpretations related to water and energy fluxes at Earth's surface.

Table 6
Overview of the Type of PTFs Used in a Set of LSMs

Model	Key references	Water movement approach	hydraulic model/function	PTF for hydraulic model/function	Additional point PTF
CABLE Community Atmosphere Biosphere Land Exchange model	Kowalczyk et al. (2006) and Wang et al. (2011)	diffusivity form of Richards' equation	Campbell (1974)	Clapp & Hornberger, 1978 lookup table according to Dirmeyer et al. (2002) modiefied from Cosby et al. (1984)[a]	PWP and FC calculated from MRC
Catchment land surface model NASA_CATCHMENT_LSM	Ducharne et al. (2000) and Koster et al. (2000)	Bucket model	Campbell (1974)		PWP and FC calculated from MRC
Community land model CLM 4.5	Oleson et al. (2013) and Lawrence et al. (2011)	Richards equation (form is not provided)	Campbell (1974)	use of various PTFs for Campbell (1974) model based on Clapp and Hornberger (1978) and Cosby et al. (1984); for organic soils all parameters are modified	not reported
Leaf/OLAM Version 4	Lee (1992), Lee and Pielke (1992), McCumber and Pielke (1981), and Walko et al. (2000)	diffusivity form of Richards equation	Campbell (1974) or van Genuchten (1980)	Clapp and Hornberger (1978) or look-up table for 12 soil classes for van Genuchten (1980) (source of parameters not provided)	PWP and FC calculated from MRC
JULES Version 4.6	Best et al. (2011)	diffusivity form of Richards equation	Brooks-Corey (1964) or modified van Genuchten (1980)[b]	modified Cosby et al. (1984)[c]	PWP and FC calculated from MRC
JSBACH version 3.0	Roeckner et al. (2003)	diffusivity form of Richards equation	modified van Genuchten according to Disse (1995) for HCC and Campbell (1974) for diffusivity	use of various PTF for Campbell (1974) model based on Clapp and Hornberger (1978), Patterson (1990), Letts et al. (2000), Beringer (2001), and Williams and Ahuja (2003)	RZWHC according to Kleidon (2004) PWP according to Hagemann (2002) Zr according to Dunne and Willmott (1996)
Noah-MP version 3.0	Niu et al. (2011)	diffusivity form of Richards equation	Campbell (1974) or van Genuchten (1980)	Cosby et al. (1984) for Campbell (1974), van Genuchten (1980) parameters derived from Cosby parameters	PWP and FC calculated from MRC
MPI-HM	Hagemann and Dümenil Gates (2003) and Stacke and Hagemann (2012)	bucket model	none	soil water capacity and wilting point calculated from different unspecified sources	regression based calculated from additional data (e.g., maximum rooting depth data set)
ORCHIDEE rev3959	Ducharne (2016)	Fokker-Planck equation	van Genuchten (1980)	Carsel and Parrish (1988) look up table for 12 USDA classes	PWP and FC calculated from MRC
SSiB Simplified Simple Biosphere model (SSiB) Version 3.5	Xue et al. (1991), Zhan, Xue, and Collatz (2003), and Sun and Xue (2001)	bucket model	extended Campbell (1974) for partially frozen soils	Clapp and Hornberger (1978) for WR and modified Clapp and Hornberger (1978) according to Jame and Norum (1980) for HC	PWP according to Xue et al. (1991)
SURFEXv8.0 version 8.0	Masson et al. (2013)	force-restore approach or diffusivity form of Richards equation	Campbell (1974)	continuous PTF according to Decharme et al. (2011) developed from Clapp and Hornberger (1978) and Cosby et al. (1984) database	either PWP and FC calculated from MRC or by PTF based on Clapp and Hornberger (1978) database

Note. FC = field capacity, PWP = permanent wilting point, WRC = water retention curve, HCC = hydraulic conductivity curve, RZWHC = root zone water holding capacity, Zr = rooting depth. [a]Revised version used for SMAP L4_SM (https://smap.jpl.nasa.gov/data/) used van Genuchten parameters according to Wösten et al. (2001). [b]Hereby van Genuchten function was coupled to Brooks-Corey parameters. [c]Reformulated to match metric units.

For a set of widely used LSMs (current as well as some historical) the hydraulic parameterization (Brooks & Corey, 1964; Campbell, 1974; van Genuchten, 1980) and PTFs used are listed in Table 6. Additionally, information of specific point PTFs for the estimation of auxiliary parameters such as permanent wilting point and field capacity are provided. Most models rely on the Campbell (1974) parameterization of the water retention and hydraulic conductivity functions. For those models the PTF of Clapp and Hornberger (1978) and Cosby et al. (1984) are mainly used, whereby Clapp and Hornberger (1978) is a so called class transfer function for 11 USDA textural classes, whereas the PTF by Cosby et al. (1984) refers to a continuous PTF using textural information (percent sand, silt, and clay) as inputs. The NASA "catchment" LSM model in comparison used the Campbell (1974) parameterization in combination with a look-up table of Dirmeyer et al. (2002), which is based on the PTF of Cosby et al. (1984). The Simplified Simple Biosphere model (SSiB), on the other hand, uses a Clapp and Hornberger (1978) PTF for the water retention curve and a modified Clapp and Hornberger (1978) PTF according to Jame and Norum (1980) for the hydraulic conductivity function. Here it has to be noted that a mixture of soil hydraulic parameterization or the use of different sets of PTFs for the water retention and the hydraulic conductivity functions may cause inconsistencies.

Brooks and Corey's (1964) soil hydraulic parameterization is only used in the JULES LSM, although note that apart from the residual moisture content it is the same as Campbell's model. In some cases, the LSM can handle different soil hydraulic parameterizations such as Campbell (1974) or van Genuchten (1980), for example, JSBach, LEAF-OLAM, NOAH-MP, or JULES, whereby JSBach is based on a modified form of the van Genuchten model according to Disse (1995). The only LSM that uses van Genuchten (1980) only is the ORCHIDEE model (see Table 6).

4.2. PTFs of Solute Transport
4.2.1. Solute Transport Processes
At present there exists a multitude of solute transport models available for a range of applications (Köhne et al., 2009a, 2009b). Among the most commonly used are the convection dispersion equation (Scheidegger, 1954) and its conceptual counterpart, the stochastic convective equation (Simmons, 1982), and two-domain models like the mobile-immobile model (Coats & Smith, 1964), dual-models like, for example, DUAL (Gerke & Van Genuchten, 1993) and MACRO (Larsbo et al., 2005). The amount of different model concepts might indicate that there is still little consensus for a parsimonious, unified solute transport model. Model choice, however, can lead to very different solute transport predictions (see, e.g., Vanderborght & Vereecken, 2007a).

Attempts to develop solute transport PTFs have, so far, been mostly restricted to small, local data sets and specific models. They have therefore very limited value for practical applications. Commonly, these PTFs aim at predicting parameters for the convective dispersive equation (Perfect et al., 2002) and the mobile-immobile model (e.g., Goncalves et al., 2001; Shaw et al., 2000). Examples for PTFs for more complex models like in Moeys et al. (2012) are rare. An extensive overview on PTF approaches for solute transport related model parameters is given in the review of Minasny and Perfect (2004).

It is difficult to relate and compare the physical meaning of specific model parameters. We instead give an overview on the state of knowledge of important factors that are influencing solute transport properties of soils using a basic, generalized transfer function to illustrate the relation between soil properties and basic solute transport features. Solute transport may be described by partitioning any respective transfer function $f(x,t)$ into four components so that

$$f(x, t) = f_{m1}(f_{m2} + f_{mh} + f_{dp}),$$ (18)

where x (L) and t (T) are space and time coordinate, f_{m1} ($L^{-1} T^{-1}$) and f_{m2} (−) are transfer functions associated with the first and second spatiotemporal moments describing the solute transport process and f_{mh} (−) stands for the transfer function for the third and all higher spatiotemporal moments. The last term f_{dp} (−) represents a source and sink term for the considered solute. The transfer functions terms were furthermore split up into contributions of water flow field with which the solutes are transported and biogeochemical interactions such as solute adsorption or chemical reactions.

4.2.2. Hydraulic Parameterization of Solute Transport
Hydraulic properties affect all four transfer-function terms. The water flow velocity, which is associated with the first transfer function f_{m1}, depends on the flow rate and the effective conducting porosity, which in turn

depends on the water saturation. Approaches for estimating the effective conducting porosity are presented in Shaw et al. (2000) and Goncalves et al. (2001). A more comprehensive PTF for the effective water filled porosity would need to include PTFs for the soil hydraulic properties as the fraction of water conducting porosity is influenced by the water saturation state of the soil.

The transfer function related to the second moment, f_{m2}, describes the average spread of a solute plume and is expressed as parameters for the molecular diffusion and the hydrodynamic dispersion, which are often combined, at least in the models that are based on the convection-dispersion equation. Several models for estimating the molecular diffusions are available of which an overview is presented in Minasny and Perfect (2004). These models estimate the diffusion coefficient from the water content, occasionally in combination with the soil porosity. Some models include a tortuosity factor instead of the porosity. It may, for example, be estimated from the water retention curve as in the model published by Moldrup et al. (2003).

The solute spread is sometimes also described as the dispersivity or, in model-independent studies, as the apparent dispersivity (Vanderborght et al., 2001). Two recent comprehensive metastudies on dispersivities have been published by Vanderborght and Vereecken (2007a, $n = 635$) and Koestel et al. (2012, $n = 733$). These studies have demonstrated that the dispersivity is positively correlated with the lateral and longitudinal scale of the experiment, the flow rate, and the clay content and negatively with the sand content, and the geometric mean particle size (Koestel et al., 2012). It is likewise known from the same study that dispersivities are smaller for repacked than for undisturbed soil by 1 order of magnitude. It follows that soil macrostructure contributes strongly to the spreading of a solute plume, provided that it is present. Vanderborght and Vereecken (2007b) found that only a quarter of the observed variance in dispersivities could be explained by a combination of the scale of the experiment, transport distance, and an interaction term of scale and soil texture. Estimating the dispersivity with larger precision accuracy hence requires including more or more suitable predictors or both.

The term "preferential flow" is well established in soil science to address transport features related to the third and all higher spatiotemporal moments of a solute plume, f_{mh}. Preferential flow is associated with localized flow paths (Hendrickx & Flury, 2001) and solute breakthrough curves with early tracer arrival times and long tailings (Brusseau & Rao, 1990). The strength of preferential flow is therefore crucial for estimating the fate of solute in the subsurface: a solute may reside for long times in a soil that is exhibiting preferential flow if it is located distantly from the preferred transport paths. But it may be flushed out quickly when located within or close to them. Ideally, the strength of preferential flow would be estimated from 2-D breakthrough data or 3-D data of a solute displacement. A suitable candidate is the dilution index-based reactor ratio (Kitanidis, 1994), which directly quantifies all higher moment features of a solute transport. For now, the respective measurement techniques for inferring to the reactor ratio, for example, by using multicompartment samplers (de Rooij et al., 2006) or X-ray imaging (Koestel & Larsbo, 2014), are still under development and there is only a minimum amount of respective data collected yet. Numerical simulations (e.g., Knudby & Carrera 2005) and metastudies (Koestel et al., 2011) suggest, however, that the arrival time of the first 5% of an applied inert solute relative to the average solute arrival time is a suitable indicator for the strength of preferential flow that can be derived from breakthrough curve data. In this respect, Koestel and Jorda (2014, $n = 558$) found that the relative 5% arrival time was strongly dependent on the clay content in a nonlinear fashion. Strong preferential flow was observed for clay contents above 10% but not below this threshold, which roughly coincides with the clay content necessary to form stable soil macrostructures (Horn et al., 2006). Furthermore, the water saturation is fundamental for estimating preferential flow strengths (Jarvis, 2007; Larsbo et al., 2014; Paradelo et al., 2016). Larger water saturations are by trend related to stronger preferential flow, indicating that macropore flow is the dominating preferential flow mechanism. Other important predictors in the study of Koestel and Jorda (2014) were the ratio of clay content and organic carbon, the lateral scale, air entrapment, and the flow rate, which all were found to be positively correlated to the strength of preferential transport. Furthermore, strong correlations between strength of preferential flow and bulk density (Koestel et al., 2013, positive correlation) and soil organic carbon content (Larsbo et al., 2016, negative correlation) have been reported. However, these studies also show that bulk density and soil organic carbon were strongly correlated with the water saturation that had established under a defined flow rate, which in turn may have been responsible for triggering preferential flow. Recent imaging-based studies have also investigated predictors for the relative 5% arrival time: Katuwal et al. (2015) found that the X-ray-derived density of the soil matrix was highly correlated to the strength of preferential flow. Karup et al. (2016) were able to estimate the arrival

times for a range of the applied solute (5% to 50%) from the clay and silt contents for 193 leaching experiments of predominantly sandy, Danish soils. Their study, however, also suggests that the correlation between the soil fine texture and the solute arrival time breaks down for the samples with clay contents of more than 10 to 15%.

Still, the validity of the above findings for relative 5% arrival times refer to transport processes observed at transient or steady state water flow, that is, for macropore flow or funnel flow (Hendrickx & Flury, 2001). They are not covering preferential flow and transport phenomena that establish predominantly under transient conditions, namely, due to flow instabilities and water repellency (Ritsema & Dekker, 1996). The former is associated with sandy soils, whereas the latter may occur at all soil textures but is induced by organic matter or biological activity (Lichner et al., 2013). We are not aware of PTF approaches aiming at predicting the strength of preferential flow for these two flow mechanisms.

Finally, the solute decay and production is also related to the water flow field because it defines the residence times of the solute in biogeochemical domains with specific conditions, which in turn may also be altered by the water flow field and the solutes that are transported with it (Jarvis, 2007).

4.2.3. Geochemical and Biogeochemical Parameterization of Solute Transport

Biogeochemical interactions influence all four transfer function terms, albeit they may be perceived as being mostly associated with the first moment of a solute transport plume and decay or production of the solute. The first moment of a solute transport plume, represented by f_{m1}, is often modeled as the product of water flow velocity and a retardation factor, which expresses the average transport velocity differences between water and solute transport, which are caused by the biogeochemical interactions. It is often inferred from adsorption isotherms measured in batch studies, that is, by creating a slurry out of sieved soil and an aqueous solution that is shaken for a defined amount of time to bring soil particles and solute into intimate contact. There are several studies on PTFs for adsorption isotherm parameters published. These mostly concern contaminants such as heavy metals (e.g., Horn et al., 2006) or excess pesticides (e.g., Kodesova et al., 2011; Moeys et al., 2011) and fertilizers (e.g., Achat et al., 2016). Virtually all these studies include the soil organic carbon content as a predictor. In fact, it is common to normalize adsorption coefficients of pesticides to the soil organic carbon content of the soil as it is argued that it provided the majority of pesticide adsorption sites. There is, however, recent evidence that such a normalization is not always justified (Jarvis et al., 2013), albeit that the soil organic carbon may still be crucial for estimating adsorption coefficients. Other common predictors for PTFs of adsorption properties are the pH and the clay content. Sometimes, also an ion exchange capacity is included. Otherwise, substance specific additional soil properties are occasionally considered, such as aluminium and calcium oxide contents in the case of phosphorus (Achat et al., 2016). A systematic overview of existing respective publications is not within the scope of this study but more publications featuring PTFs for adsorption isotherms are listed in Minasny and Perfect (2004). A caveat for measuring adsorption isotherm in batch experiments is that the lab conditions differ considerably from the conditions encountered in the field (Vereecken et al., 2011). The latter authors found in their metaanalysis that there was a systematic mismatch between batch and leaching experiment derived adsorption coefficients that was dependent on the water saturation, the packing procedure used for the soil columns, and the water flow velocity. It is noted that in the case of nonlinear adsorption and kinetic adsorption, also the higher moments of a solute plume are influenced, that is, f_{m2} and f_{mh}. Nkedi-Kizza et al. (1984) demonstrate this for the latter case by showing that the equations of the mobile-immobile model for chemical and physical nonequilibrium are mathematically identical.

Establishing PTF approaches for estimating solute degradation or production rates in soil, f_{dp}, is very difficult since f_{dp} depends strongly on the presence of microbial communities and their activity. Among the few studies that have so far tried to take on this challenge, von Gotz and Richter (1999) identified initial microbial biomass to predict the degradation of the pesticide bentazone. Fenner et al. (2007) developed a quadratic regression relationship using soil pH, organic carbon and sand content, annual mean temperature, and the sampling depth and sample height of the respective soil sample to appraise atrazine persistence in soil. Some of the listed predictors are proxies for atrazine adsorption, that is, its availability to microorganisms. Among these predictors are the organic carbon and sand contents. Other predictors, like the soil pH and the temperature, are estimators for the quality of habitats for soil microorganisms. Ghafoor et al. (2011) extended a similar approach to metadata on the degradation of 16 pesticides that had been collated from peer-reviewed literature. They found that the organic matter and clay content as well as the soil depth,

the microbial biomass, the water content, and the adsorption isotherm of the respective soil-pesticide combination were important estimators for pesticide degradation.

From the above it is obvious that PTFs for solute transport parameters require knowledge of the hydraulic properties. In most practical cases this entails that PTFs for solute transport properties must be built on top or together with PTFs for hydraulic properties. Approaches following the ones presented in Shaw et al. (2000) and Goncalves et al. (2001) seem therefore reasonable, as long as they are built upon a larger database. Furthermore, the scale of the transport process must be considered in a respective PTF. A similar conclusion had already been reached by earlier, less comprehensive studies (Kolenbrander, 1970; Perfect, 2003). In this context Vanderborght and Vereecken (2007b) notably suggest that the dispersivities are lognormally distributed in space, which may help creating upscaling relationships for solute transport PTFs. Also, predictors for soil macrostructure are important. These include the clay content and the soil organic carbon content. Approaches that also include biological activity as a structure forming agent like the one of Jarvis et al. (2009) are considered promising. The above statements concern inert solute transport. PTFs for reactive solute transport require a plethora of additional information that are solute specific and concern the soil chemistry (e.g., pH, redox conditions, and ion exchange capacities) and the soil microbiology and, probably, also additional soil biological features, for example, the presence or absence of specific root exudates.

The establishment of PTFs for solute transport is still work in progress, and respective approaches have barely reached a precision that is required for practical applications. Larger databases on solute transport data are needed for building and validating the PTFs. This conclusion echoes the one of Minasny and Perfect (2004), albeit some efforts toward larger respective databases have meanwhile been undertaken (see Koestel et al., 2012; Vanderborght & Vereecken, 2007b). New databases should also include data from neighboring research fields and from new measurement approaches like remote sensing, isotope analyses, omics data, imaging techniques, and critical zone observatories (Li et al., 2017).

4.3. PTFs of Heat Exchange

Thermal properties like heat capacity or thermal conductivity are needed for calculation of the thermal regime in vadose zone and Earth system models, including LSMs. Subsurface temperature profiles are in the vast majority of cases derived from the general heat diffusion equation in the vertical direction. In this way, the change in soil temperature for each soil layer, at each model time step, depends on the flow of heat into and out of that layer, the thickness of the layer, the length of the time step, and the volumetric heat capacity, C_h, of each layer (C_h is often called "apparent" heat capacity if phase changes are included; see, e.g., Cox et al., 1999). In some cases, advection of heat by water flow is considered as an extra term in the heat diffusion equation (e.g., Best et al., 2011). The flow of heat is determined by a vertical temperature gradient multiplied by a thermal conductivity, λ; this is generally referred to as Fourier's law.

Thermal properties can be measured directly in the field using heat needle probes or in the laboratory (Horton et al., 1983). However, because these methods are time consuming, especially in light of the spatial and temporal variability of these properties, equations have been established for the calculation of thermal properties that depend on soil textural, structural properties (de Vries, 1963; Farouki, 1986; Johansen, 1975), and on soil water content and its phase (liquid or frozen). Provided that the volumetric fractions of moisture (liquid and/or frozen), minerals, and organic matter are known, the heat capacity can be calculated easily (de Vries, 1963). The estimation of thermal conductivity relies on empirical models, analogous to the hydraulic functions, and similarly, with uncertainties in parameter values (Peters-Lidard et al., 1998; Tarnawski et al., 2011). For example, the present thermal conductivity models face an uncertainty in parameterization of the shape factor of soil particles (Lu et al., 2007). To increase confidence in these models, more data of measured thermal conductivities are needed, gathered from a wider range of soil types (Stolte et al., 1996), as the equations are based on a relatively small number of measured data, in particular, compared to the number of samples available for derivation of the hydraulic PTFs (see Table 4). Recent advances to improve vadose zone and land surface models are looking to solve this uncertainty (Calvet et al., 2016; Decharme et al., 2016; Dong et al., 2015; Lu et al., 2007; Tarnawski et al., 2009).

Information on soil texture (in all cases, and in some cases quartz content, see also Calvet et al., 2016), and organic matter (in most models) is required to calculate key parameters in the equations that describe soil thermal properties as a function of soil moisture content or relative saturation. Although these are not

traditionally referred to as pedotransfer functions they can definitely be viewed as such. Soil porosity, required to calculate relative saturation, for example, can be derived from the hydraulic PTFs described in section 4.2. Below the broad types of thermospecific equations and their "PTFs" are discussed.

Soil heat capacity, C_h, is generally a linear function of the heat capacities of the soil solids (mineral and organic matter), liquid water, and ice constituents (Van Wijk & de Vries, 1963). In order to get the bulk heat capacity required in the general heat diffusion equation, these component heat capacities (see, e.g., Johansen, 1975; Van Wijk & de Vries, 1963) are weighed by their constituent volume fractions (so that the heat capacity for water is multiplied by soil moisture content). Simple thermal PTFs are used to calculate the solid C_h values (dry heat capacity) as a function of mineral and organic content. There are some subtle differences between models, relating mainly to how mineral and organic components are weighed, but overall, LSMs follow a very similar approach.

Equations and related thermal PTFs employed to calculate soil thermal conductivity, λ, are more complex and differ considerably among LSMs. Most thermal conductivity parameterizations in the LSMs are predicted based on the work of Johansen (1975) and Farouki (1986). This includes, for example, CLM, ISBA-SURFEX, and JULES. Soil moisture content is considered to be the main driver of change for changes in thermal conductivity (Sourbeer & Loheide, 2015; Subin et al., 2013). The value of λ depends nonlinearly on the way in which the best conducting mineral particles are interconnected by the less conducting water phase and separated by the poorly conducting gas phase (Koorevaar et al., 1983). In the case of λ a weighing function, generally called the "Kersten" equation (e.g., Johansen, 1975), is employed, to express this soil moisture dependency. It is a function of the degree of saturation and different Kersten functions can be used depending on the phase of water (frozen or unfrozen soil water). Generally, λ varies between a minimum (dry) and saturated λ value. The dependence on relative saturation is often logarithmic, or as a quadratic function in the case of the OLAM model (Walko et al., 2000).

Thermal PTFs are mainly required to calculate these "dry" and "saturated" thermal conductivity and in some cases to get parameters relating to the shape of the "Kersten" function. These PTFs require thermal conductivities of solid soil components (clay, silt, sand, and organic matter), water, and in some cases of air, as well as porosity or dry bulk density. In some cases the mineral solid components are split into quartz (with a high thermal conductivity around 8 W m^{-1} K^{-1}) and other minerals (2 W m^{-1} K^{-1} for quartz content >20%, 3 W m^{-1} K^{-1} for lower values). In most cases quartz fraction is directly related to sand content, but recently, PTFs for quartz fraction have been proposed in a LSM context (e.g., Calvet et al., 2016).

With soil moisture determining the temporal pattern of thermal conductivity (Sourbeer & Loheide, 2015; Subin et al., 2013), for given soil moisture conditions, thermal conductivity spatially depends to a large extent on the fraction of soil minerals presenting high thermal conductivities such as quartz, hematite, dolomite, or pyrite (Côté & Konrad, 2005). In midlatitude regions of the world, quartz is the main driver of thermal conductivity for mineral soils, and PTFs for the quartz fraction are proposed in this context (Calvet et al., 2016). In organic rich soils the soil organic matter (SOM) content is most relevant. With actual global soil grid information available for SOM, gravels, and bulk density (Hengl et al., 2014; Nachtergaele et al., 2012), application of PTFs to derive porosity and quartz fraction now allows estimation of heat properties to soils globally. The restriction to the use of these PTFs for now is for soils containing a rather large amount of organic matter as the derived PTFs are currently valid for volumetric fractions m_{sand}/m_{SOM} ratio values lower than 40 (Calvet et al., 2016).

4.4. PTFs of Biogeochemical Processes
4.4.1. Soil Carbon and Nutrient Cycling Processes
Earth system models explicitly model the movement of carbon through the Earth system. Still, there is an overall lack of spatially explicit models that properly describe soil carbon and nutrient dynamics at different spatial scales (Manzoni & Porporato, 2009). In many cases input data for land surface models cannot be measured directly such as the initial carbon pool size distribution or root water uptake and need to be estimated. Therefore, a great demand exists to expand available databases by states and parameters derived from model inversion/data assimilation or statistical approaches to develop PTFs. The estimation of model parameters using field data is a standard procedure to calibrate models to site specific conditions for soil carbon and nitrogen turnover (Bauer et al., 2012; Carvalhais et al., 2008; Klier et al., 2011; Ludwig et al., 2007),

crop growth models (Billen et al., 2009; Priesack et al., 2006), and long-term experimental data (Bauer et al., 2012; Sakurai et al., 2012). Soil carbon dynamics are typically conceptualized by multicompartment approaches where each compartment is composed of organic matter with similar chemical composition or degradability (Bricklemyer et al., 2007; Coleman et al., 1997). Nitrogen turnover is strongly related to carbon turnover, and both are often part of overall land surface and terrestrial ecosystem models (Batlle-Aguilar et al., 2011; Manzoni & Porporato, 2009; Priesack et al., 2008).

4.4.2. Soil Carbon Parameterization

The frequently applied, yet criticized, conversion factor of 1.724 (van Bemmelen, 1890) to estimate the SOM from the SOC organic carbon (SOC) content can probably be seen as one of the earliest PTFs. So also in the field of biogeochemical or ecological research many of those simple conversion rules exist and are commonly applied to convert the soft information of soil maps into hard numbers.

Against the background of climate change the exact quantification of soil carbon stocks is of paramount relevance. Carbon stocks are given as mass C per unit area, which requires the reference depth, the gravimetric SOC content of that layer, and the soil bulk density. The application of a PTF to estimate bulk density could thus lead to substantial uncertainty in the carbon stock prediction (Kobal et al., 2011; Xu et al., 2015), particularly for repeated inventories to quantify temporal changes in soil carbon stocks (Schrumpf et al., 2011). Many PTFs to estimate bulk density exist. However, their application is often limited to a specific region (Bernoux et al., 1998; De Souza et al., 2016), forest soil (Kobal et al., 2011; Nussbaum et al., 2016), or a soil taxonomic database (Benites et al., 2007; Manrique & Jones, 1991). PTF estimates of bulk density are always based on SOC content according to the model of Adams (1973) where bulk density of a soil can be modeled as mixture of organic matter and mineral component. The mineral components expressed as clay and/or silt content are identified as statistically significant explanatory variables (Bernoux et al., 1998; Benites et al., 2007; Botula et al., 2015; De Souza et al., 2016; Manrique & Jones, 1991). Also, the coarse fraction ($>$2 mm) and soil depth has predictive potential (Nussbaum et al., 2016) as well as the sum of basic cations (Benites et al., 2007) and other environmental covariates (De Souza et al., 2016). Based on clay content, silt content, soil color, and depth, Minasny et al. (2006) derived PTFs for SOC contents and bulk density and applied exponential depth functions to subsequently estimate carbon stocks for a 1,500 km^2 region in Australia. Hengl et al. (2017) gathered global-scale SOC and bulk density data, and based on classical PTFs, they also estimated SOC contents and provide estimates of global soil carbon stocks and their uncertainty.

The turnover of soil SOC represents a crucial component of the global carbon cycle and models are valuable tools to improve understanding and make predictions of carbon sequestration and soil CO_2 emission. As an example Wang et al. (2016) estimated the carbon inputs to wheat systems required to maintain current SOC stocks at global scale. They applied the PTFs of Weihermüller et al. (2013) to estimate the initial SOC model pools from the SOC and the clay content assuming equilibrium. For the RothC model (Coleman & Jenkinson, 2005) another PTF by Falloon et al. (1998) exists, which allows the estimation of the inert organic matter pool from the total SOC content, even though the authors did not explicitly refer to this as a PTF. Also, the regression equations determined by Zimmermann et al. (2007) for the estimation of the ratio between the resistant and the decomposable plant material pool of the RothC model was not referred to as a PTF by the authors. The PTFs of Falloon et al. (1998), Zimmermann et al. (2007), and Weihermüller et al. (2013) could, analogous to the continuous PTFs for the estimation of soil water retention parameters, be classified as parameter estimation PTFs since model specific parameters, which could not be measured directly, are estimated. Also, the PTFs of Moore et al. (1992) fall into to this class as the parameters of the dissolved organic carbon sorption isotherm were estimated from SOC, oxalate-extractable Al, and dithionite-extractable Fe contents.

In addition, the C saturation level or the capacity of soils to store organic C can be predicted as a function of clay and silt content by Hassink (1997) or clay content (Dexter et al., 2008). These authors postulated that carbon has to be associated with clay (and silt particles) to form microaggregates that enable the carbon to be physically stabilized.

4.4.3. Soil Mineralization and Decomposition Parameterization

Against the background of greenhouse gas emissions and nutrient cycling, nitrogen mineralization is a crucial process. Most of the coupled C and N biogeochemical models include plant N uptake, soil organic N decomposition, microbial N mineralization and immobilization, biological N fixation, and various pathways of N export (Xu-Ri & Prentice, 2008). All models follow the mass balance principle where the inorganic N

added to an ecosystem equals the cumulative changes in plant N uptake, soil N retention, and N loss through leaching and gaseous emission (Zaehle & Dalmonech, 2011). Rasiah (1995) compared existing PTFs to estimate the one and two pool sizes and related rate constants of N turnover based on total N, cation exchange capacity (CEC), SOM, and other soil properties. Rasiah and Kay (1998) developed PTFs for the estimation of N mineralization parameters in dependence of the fraction of the total volume that is air-filled and the fraction of the air-filled porosity. A PTF for the estimation of the slow N pool size in sandy arable soils based on SOC, total N, C:N ratio, and mineral fraction <20 μm was proposed by Heumann et al. (2003). A larger textural spectrum of soils located in northwestern Germany was covered by Heumann et al. (2011) where regressions of the fast and the slowly mineralizable N pool size with clay content, humus class, mean fall temperature, and total N were established following the analysis of incubation experiments. Against the background of potential NH4 leaching risk Vogeler et al. (2011) identified clay content and the CEC to be most related to the Freundlich isotherm parameters for NH4 adsorption. Glendining et al. (2011) compiled a rather extensive database on total N measurements in soil and derived regression-based PTFs for global-scale estimates. Highest predictive power for total N at global scale was detected for SOC, latitude, soil group factor, and texture class as inputs.

Phosphorus stimulates C and N mineralization and is also a limiting nutrient for plant production. Remaining P is a simple index indicating the adsorption of P in soil. Cagliari et al. (2011) derived a PTF for remaining P in dependence of pH, sum of exchangeable bases, and exchangeable Al, however, for a relatively small set of soils located in the Sao Paulo state, Brazil. Also, based on the investigation of Brazilian soils, Camargo et al. (2016) derived a PTF for the adsorption of P accounting for iron oxide content and magnetic susceptibility. Cohen et al. (2007) investigated the sorption of P in wetland soils in southeastern USA in dependence of either visible/near-infrared diffuse reflectance spectra or biogeochemical properties and found comparable predictive capacities of both approaches. Using a global data set, Achat et al. (2016) derived PTFs for inorganic P availability and identified oxides and SOC as best explanatory variables, even for nonforest soils.

Other biogeochemically relevant soil properties like CEC could be related to clay content, SOM, and pH via PTFs (Bell & van Keulen, 1995). Liao et al. (2015) additionally identified silt and sand content as having significant predictive potential for the estimation of CEC. Soil-specific surface area could also be estimated from textural composition (Whitfield & Reid, 2013) or from additional information on plastic limit, liquid limit, and free swelling index (Bayat et al., 2015).

The vast majority of PTFs developed in a biogeochemical context rely on regression equations. ANNs were applied by Cagliari et al. (2011) and Bayat et al. (2015), whereas De Souza et al. (2016) used random forests.

5. Challenges for PTFs in Earth System Science

The necessity to improve LSMs and terrestrial ecosystem models is timely, and here we look for the possibilities PTFs offer with specifically new opportunities and solutions for parameterization of biogeochemical and biotic processes.

There are two avenues to bring about this improvement: one is to translate current knowledge on environmental relationships into spatially exploitable PTFs to parameterize specific processes and the other is looking to model coefficients currently set constant, which appear to be spatially and temporally variable. So either improved knowledge and data availability on environmental relationships can be exploited, as we present for vegetation parameters in section 5.1, and biodiversity and biotic processes in section 5.2, or currently fixed parameters or proxies might be improved with PTFs, as demonstrated for the Q10 temperature coefficient in section 5.4. This approach of soil-based spatial (geographical) weighting can be extended to all equations/relationships that need parameterization, and that could benefit from existing fine-grained soil information to make spatial extrapolations. The availability of soil data globally that provides information up to 250 m resolution of texture, bulk density, and organic matter along the depth of soil profiles (Hengl et al., 2017) offers unseen opportunities for the development and application of PTFs in Earth system science. The ambition here is to show how PTFs can be improved to better parameterize soil processes by better quantifying rate parameters in biogeochemical processes related to soil carbon (e.g., Q10 and first-order rate constants), nitrogen (e.g., mineralization constants), and biotic and biodiversity-related processes. The latter has been largely neglected as there is a rising awareness of the different soil biotic processes and their role in the C, N, and general nutrient cycling in terrestrial ecosystems (Filser, et al., 2016).

With that exercise, challenges arise for increasing understanding and modeling of climate change influences on soil processes, biota, and biodiversity controls to the biogeochemical processes and furthermore impacts of land use changes and (adaptation) management. These knowledge gaps are recently addressed in many specific studies and are topical in Earth system science in general. Yet integration of the knowledge of different disciplines remains challenging.

5.1. Applying Soil Information to Improve Vegetation Parameter Estimation

5.1.1. Replacing Plant Functional Types by PTF-Based Vegetation Parameterization

LSMs usually classify plant species into plant functional types, within which all parameters are identical. This abstraction is necessary for simulating large geographical scales. However, current LSMs have only about 3–12 plant functional types, and hence, they typically ignore most biodiversity within a simulation grid (Sato et al., 2015). This oversimplification can lead to LSMs overestimating the strength of some climate responses, since it neglects adaptation capacity. Negative effects on vegetation due to climate change can be mitigated by increases in those species best adapted to the new conditions (Purves & Pacala, 2008). Adaptation of plants and plant communities to the soil and functional interactions between soils and plants are not currently included in plant functional types.

On the global scale, basic plant functional classifications capture an important fraction of trait variation to represent functional diversity (Kattge et al., 2011). This assumption is implicit in today's dynamic global vegetation models, used to assess the response of ecosystem processes and composition to CO_2 and climate changes. Owing to computational constraints and lack of detailed information, these models have been developed to represent the functional diversity of 4,300,000 documented plant species on Earth with a small number (5–20) of basic plant functional types. This approach has been successful, so far, but limitations are becoming obvious and challenge the use of such models in a prognostic mode, for example, in the context of Earth system models (Lavorel et al., 2008; McMahon et al., 2011). The plant functional types capture a substantial fraction of the observed variation, but for several traits most variation occurs within plant functional types. In the context of vegetation models these traits would better be represented by state variables rather than fixed parameter values (Kattge et al., 2011).

There are new approaches in Earth system modeling to better account for the observed variability: suggesting more detailed plant functional types, modeling variability within plant functional types, or replacing plant functional types by continuous trait spectra. The global plant trait database TRY (www.try-db.org) provides information about ecological strategy type, phenology, morphology traits, and habitat characteristics. It contains a large data set that gives an important input for the definition of new, more detailed plant functional types, and in particular for independently parameterizing the plant traits (Kattge et al., 2011). Defining trait ranges and state variables based on geographic (biomes/bioclimates/plant geographic regions) and soil conditions would allow for a plant functional type-based parameterization of plant traits in the dynamic vegetation models. This is a more than obvious step since soil conditions of moisture retention and nutrient provision to plants crucially determine the essentially used traits for the vegetation models (specific leaf area, foliage nitrogen concentration). Some recent models represent trait ranges as state variables along environmental gradients rather than as fixed parameter values. The dynamic global vegetation model O-CN (Zaehle & Friend, 2010) is an example of a land surface model development toward such a new generation of vegetation models, also the nitrogen carbon interaction model (Esser et al., 2011), or in combination with an optimality approach toward vegetation water use the vegetation optimality model (Schymanski et al., 2008). The earlier discussed rooting depth and density (section 4.1.2) could in the same way benefit from this proposed combination of information of soil conditions and plant traits—instead of the rough plant functional type classification. Rooting depth (or more exactly, maximum water extraction depth) is among the most influential plant traits in global vegetation models, yet we have estimates for only about 0.05% of the vascular plant species—as measurements are difficult and laborious. However, many aboveground traits correlate with belowground traits (Kerkhoff et al., 2006), so the data in TRY could be indicative about belowground traits. Especially, in combination with the here proposed coupling to soil information. Furthermore, it is important to understand the feedbacks between environmental/climate change and subsequent responses of vegetation (Engelbrecht et al., 2007; Hare et al., 2003; Hartnett et al., 2013; Rodriguez et al., 2008; von Arx et al., 2012) and crops (Jones, 2007); many of these feedbacks come together in the notion of coevolution (Pelletier et al., 2013). Coevolution may depend on the survival strategy of a particular species and ask for different

approaches regarding PTF development. For example, it may be beneficial for perennial species to adjust their root length density to a certain soil type in order to encertain optimal water supply defined by the water storage capacity over the root zone, which may lead to a reduced expression of soil variability. While annual plants may have invested more in seed dispersal and be more prone to senescence when soil moisture becomes depleted, and thus express soil varaibility in full. Coevolved soil-plant relationships have been shown for photosynthetic traits, soil pH, available phosphorus, and ratio of precipitation to potential evapo-transpiration (Maire et al., 2015).

5.1.2. PTFs for Parameterization of Vegetation Water Content

Not only the application of plant functional types, but also the estimation of crucial vegetation para-meters associated with these plant functional types, like in section 4.1.2 described root density, but also the vegetation water content (VWC), and other functional traits for processes of transpiration, transforma-tion, and ecosystem production, could be strongly improved by entering soil information in their estima-tion. Nowadays, these parameters are mostly derived from satellite measured leaf area index (LAI) and normalized difference vegetation index (NDVI), combined with Moderate Resolution Imaging Spectroradiometer (MODIS) land cover data (Jackson et al., 2011). VWC is often even entered as a con-stant in functions of land surface models. To obtain a relevant estimation of these vegetation parameters, applicable in model estimation of gas exchanges and ecosystem production, basic soil properties like water and organic carbon content obviously bring in additional information to just the canopy cover esti-mation in NDVI/LAI. Especially, VWC should be improved in its parameterization, for which current meth-ods use NDVI and MODIS land cover derived stem factor (an estimate of the peak amount of water residing in the stems).

$$VWC = \left(1.9134 \times NDVI^2 - 0.3215 \times NDVI\right) + stem\ factor \times \frac{NDVI_{max} - NDVI_{min}}{1 - NDVI_{min}}. \qquad (19)$$
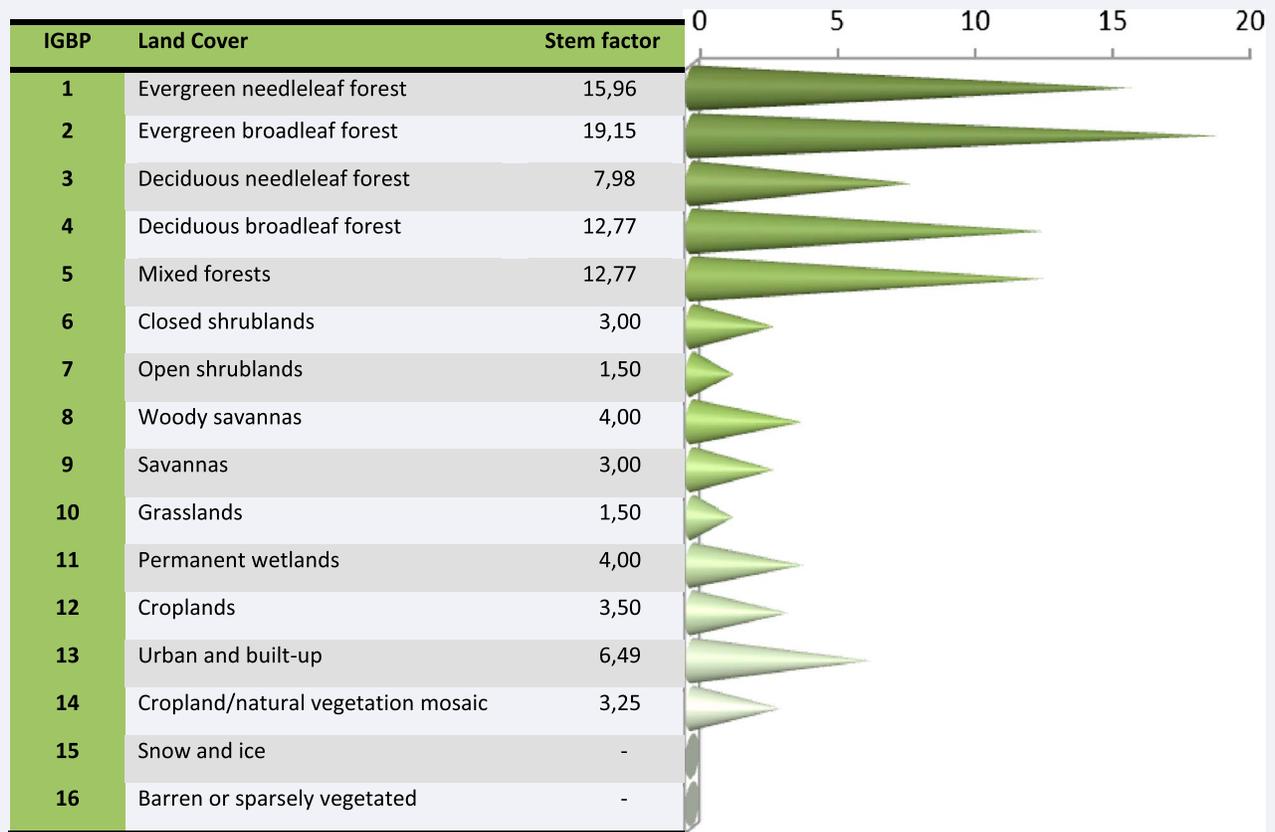
With NDVI ranges between -1.0 and 1.0, the stem factor (with ranges up to 20) has a strong impact on VWC.

The currently used stem factor estimates in Table 7 are based on roughly estimated average values for LAI and tree height for a given biome in Hunt et al. (1996). Obviously, water content of vegetation can vary strongly within biomes, depending on soil water and organic matter content, for which improved functions might be derived based on improved estimations of these basic soil properties. Here we argue that there is sufficient data available to derive relationships between biomass production, vegetation water content (based on dry weight methods biomass measurement), and soil conditions over the differ-ent biomes, land use, and vegetation types. Globally derived PTFs for these soil parameters of water and organic matter content might thus enable to obtain improved estimations of VWC, adding to the satellite-derived information.

Where the satellite derived indices can detect "greenness" and thus directly indicate canopy density (vegetation parameters of LAI and phenology), they are also generally considered robust proxies for other ecosystem elements of vegetation dynamics and production (Jackson et al., 2011). These applications prove successful for (homogeneous) rangelands and crop systems. For more heterogeneous (layered/structured) vegetation types like forests, these measures prove little accurate for detection of production or vegetation water content detection. NDVI is definitely useful in estimating photosynthetic activity for model purposes (prediction oriented), possibly even for estimating production to a certain extent (descriptive), but not as proxy for VWCs in processes like transpiration and other water flow/retention pro-cesses with important implications to model predictions of ecosystem impacts of climate changes. For instance, where a recent study indicates a general strong correlation between evapotranspiration and NDVI for an urban environment with different vegetation types (Nouri et al., 2014), which they consider promising for global applications, they document the correlation to disappear under water stress condi-tions. So even though one could argue that current NDVI-derived VWC estimates are sufficiently differen-tiated for global model purposes, when applied in the perspective of climate change impacts, the estimation of effects of drought stress to vegetation becomes essential. The improved understanding of responses of plants to drought stress ruled by water potentials in soils stems, strongly depending on soil type/properties (Malavasi et al., 2016), offers crucial information to feed such models, possible through improved parameterization of VWC with PTFs.

Table 7
Stem factors for different MODIS IGBP land cover types

IGBP	Land Cover	Stem factor
1	Evergreen needleleaf forest	15,96
2	Evergreen broadleaf forest	19,15
3	Deciduous needleleaf forest	7,98
4	Deciduous broadleaf forest	12,77
5	Mixed forests	12,77
6	Closed shrublands	3,00
7	Open shrublands	1,50
8	Woody savannas	4,00
9	Savannas	3,00
10	Grasslands	1,50
11	Permanent wetlands	4,00
12	Croplands	3,50
13	Urban and built-up	6,49
14	Cropland/natural vegetation mosaic	3,25
15	Snow and ice	-
16	Barren or sparsely vegetated	-

5.2. Biotic Processes and Biodiversity in Biogeochemical Processes

5.2.1. Biotic Process Parameterization

Most of the important soil physical properties can be predicted by PTFs, but little work has been done on soil biological properties (McBratney et al., 2002). Fundamental biotic processes like photosynthesis can be predicted based on vapor pressure deficit and soil moisture content but without "accurate PTF" at hand (Rogers et al., 2016). Filser et al. (2016) suggest that inclusion of soil biota activity (plant residue consumption and bioturbation altering the formation, depth, hydraulic properties, and physical heterogeneity of soils) can strongly affect the predictive outcome of SOM models, yet this soil biotic activity can on its turn be predicted from soil properties (Bardgett & van der Putten, 2014). And, of course, there are many more links, for which it is important here to draw the perspective of integration in Earth system models.

Approaches to model temporal changes of soil structure, biotic activity, and root growth are relatively rare (e.g., Leij et al., 2002; Stamati et al., 2013). There are still relatively few models of interactions between physical and biological processes (e.g., Laudone et al., 2011; Šimůnek et al., 2009; Tartakovsky et al., 2009). Recent advances in measurement technologies have provided new insights about the role of soil biota and biodiversity on soil and crop processes generating new knowledge and opening new perspectives for their mathematical description.

Recent models describe both the mechanical processes of soil biota (Barot et al., 2007; Blanchart et al., 2009), food web properties (de Vries et al., 2013) and biotic compartment contributions to C and N cycles in soil (Huang et al., 2010; Li et al., 2017). Over the past decade, major advances have been made in incorporating plant and soil N cycling processes in the terrestrial ecosystem models. The TEMs nowadays include plant and soil nutrient cycling processes and N constraints on the C cycle,

as reviewed by Thomas et al. (2015). These authors describe the different methods of how N limitations are currently implemented and evaluated in TEMs used in Earth system models. They point out some challenges to the development and inserting of N/C-coupled TEMs in Earth system models, challenges to temporal and spatial scales, to the lack of observation data for their systematic evaluation, and challenges associated with mechanistic representation of plant controls on N availability and turnover, including N fixation and organic matter decomposition processes. Similarly, Zaehle and Dalmonech (2011), observe that C-N coupled terrestrial ecosystem models use different approaches to parameterize complex, nonlinear C-N processes and linkages. This limits the validation of predicted N cycling impacts on terrestrial C budgets, and due to the scarcity of relevant observed data sets to evaluate the performance of coupled C-N models, the validity of C-N models is inferred from their ability to reproduce local, regional, and global features of the C cycle and comparing them with associated C-only model versions (Bonan & Levis, 2010; Huang et al., 2016; Zaehle & Friend, 2010). Global models now correct for the nitrogen limitation and general plant controls to biogeochemical processes (Niu et al., 2016; Thomas et al., 2013) but still lack many biotic controls that govern the (nonlinearity of) C-N processes and linkages. Same way, existing ecosystem productivity models (based on C models) show differences >25% even for temperate prairie grassland regions where much effort to construct such models is concentrated (Huang et al., 2016). Nonlinear models integrating pore volume and water content to describe microorganism growth and formation of organic matter (and conservation in oxygen-less microzones) are recently developed (Vasilyeva et al., 2016), opening alleys for PTF development. Yet much improvement of knowledge bases in this domain is still possible.

5.2.2. PTFs at Hand and Potential Development

With more accurate estimation of the soil variables of soil pH and N, and especially the knowledge on soil temperature and water retention parameters, the development of better performing models is achievable using PTFs to estimate the soil microbial biodiversity and C-N processes. Indeed, the relative importance analysis of regressors supported the finding that soil water availability variables are the main drivers of C_{Mic}, alone accounting for more than 80% of the explained variance (Serna-Chavez et al. 2013). Still, study in Mediterranean semiarid conditions revealed higher microbial biomass under dry conditions than under mesic conditions (Sherman et al., 2012). So here again specific relationships for biomes need to be identified. Estimates of the global storage of soil microbial biomass C and N in 0–30 cm and 0–100 cm soil profiles are derived, based on specific relationships for the major biomes (boreal forest, temperate coniferous forest, temperate broadleaf forest, tropical and subtropical forests, mixed forest, grassland, shrub, tundra, desert, natural wetland, cropland, and pasture), at 0.05° by 0.5° spatial resolution (Xu et al., 2013).

Major soil microbial and mycorrhizal activity of heterotrophic respiration and nutrient mineralization are strongly affected by soil moisture content (Zhao et al., 2016). Both hydrologic variability and its effects on soil moisture (Rodriguez-Iturbe & Porporato 2004) transferred on decomposer activity need to be accounted for in mathematical models of soil C and nutrient cycling. The effects of soil moisture on decomposer activity and decomposition rates are typically described by kinetic rate correction functions defined at the whole decomposer-community level. When soil moisture decreases, so does decomposer activity, with strong differences in the responses of individual decomposer groups to moisture availability (e.g., bacteria are typically more sensitive than fungi to water stress). The shape of the soil-moisture response and the threshold soil moisture at which decomposer activity effectively ceases (the water-stress threshold) varies across biomes and soil types and thus the responses of organism types (e.g., soil fauna, bacteria, and fungi) (Freckman, 1986). Manzoni et al. (2012) showed that responses of decomposers at the community level are different in soils and surface litter but similar across biomes and climates. This results in a nearly constant soil-moisture threshold corresponding to the point when biological activity ceases, at a water potential threshold comparable to the soil moisture value where solute diffusion becomes strongly inhibited in soil, while in litter it is dehydration rather than diffusion that likely limits biological activity around the stress point. Because of these intrinsic constraints and lack of adaptation to different hydroclimatic regimes, changes in rainfall patterns (primary drivers of the soil moisture balance) may have dramatic impacts on soil carbon and nutrient cycling (Manzoni et al., 2012). The kind of relationships between activity of organism groups (bacteria and fungi) derived in these studies allows for deduction of relevant PTFs for soil functions over larger territories and biomes.
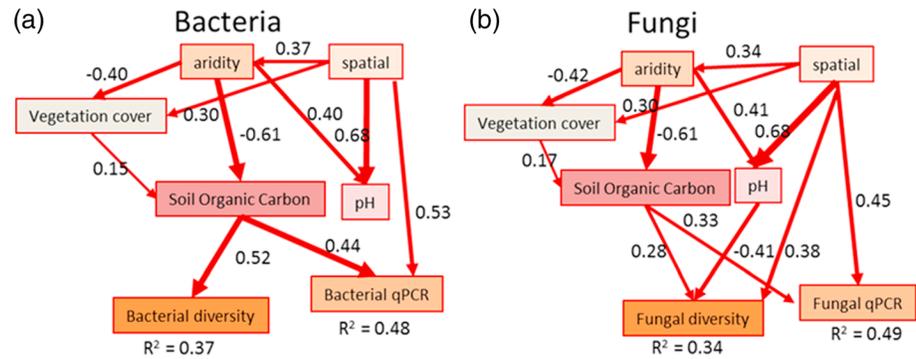
Figure 6. Structural equation models fitted to the diversity and abundance (qPCR) of (a) soil bacteria and (b) fungi. Measured relationships between soil biodiversity and soil organic C content, pH, aridity (measure of soil moisture based on soil water retention), total plant cover, latitude, and longitude indicate the potential development of inference based on PTFs for Earth system models (adapted from Maestre et al., 2015).

Current terrestrial ecosystem models in Earth system models are mostly based on multivariate regressions with more robust climatic variables as independent variables. This also applies to the models proposed to predict soil biodiversity. Regional to global soil biodiversity mapping is mostly based on microbial diversity, which is proved to correspond to microbial mass estimates (Fierer et al., 2009), based on the microbial carbon C_{Mic} (Serna-Chavez et al., 2013). Soil profile C_{Mic} seems best explained by a combination of precipitation, evapotranspiration, temperature, soil pH, and total nitrogen. The model presents a typical combination of classic large-scale distribution modeling elements of coarse-grained climate variables (precipitation and temperature) that are robust in predicting patterns at global to continental scales.

The biodiversity ecosystem function relationship is underlying most actual soil biodiversity study (Bender et al., 2016; Bradford et al., 2014; He et al., 2009; Jing et al., 2015; Ramirez et al., 2015), and some authors even go as far as stating that soil quality is primarily determined by soil biodiversity (Gardi et al., 2008, 2013). The premise to these assumptions is that more diverse soil communities exhibit higher functional diversity and thus play a stronger role in more soil ecosystem functions, directly or indirectly (Eisenhauer et al., 2012; Lefcheck et al., 2015). This straightforward assumption is backed up by some ongoing modeling attempts for specific groups like earthworms (Zaller et al., 2011) or soil microbial and fungal diversity (Fierer et al., 2009). Soil bacteria and fungi are largely responsible for key ecosystem services, including soil fertility and climate regulation, yet our understanding of these relationships—especially in the light of changing environmental conditions—is still in its infancy. Soil microbial diversity is documented to be related to soil moisture, C/N ratios, and pH (Fierer & Jackson, 2006; Manzoni et al., 2012). From a global dryland study, recent studies report that contrary to previously reported global-scale results (Fierer & Jackson, 2006), soil pH was not a major driver of bacterial diversity, in contrast to soil fungal diversity (Figure 6, Maestre et al., 2015).

The application of PTFs to validate the reported relationships between soil microbial biodiversity and soil parameters (SOC, soil moisture, pH, and soil/root depth) specified per biome might enhance our understanding and predictive capacities to soil respiration and nutrient mineralization under changing environmental conditions. Relationships between soil fungal diversity and vegetation (crops and trees) resistance to changing climatic conditions, together with water holding capacity of both soil and vegetation, are recently acknowledged and might lead to more structured application of this knowledge—through the construction of PTFs—in terrestrial ecosystem models and LSMs.

5.3. Heterogeneity of Soils, Hot Spots of Soil Processes, and Functions

More and more evidence has been gathered recently for the existence of hot spots for biogeochemical processes like denitrification (McClain et al., 2003; Palta et al., 2014), carbon sequestration (Khosravi Moshizi et al., 2015; Timilsina et al., 2013), and also of biodiversity (Mittermeier et al., 2011). These hot spots are mostly associated to distinct "critical interfaces" as interactive boundaries between zones of distinct hydrological, biogeochemical, and topographic properties. Reaction rates at these interfaces are often orders of magnitude

higher than the rest of the domain. These discrete interfaces—like hyporheic zones or biofilms—can in some cases dominate reactions within an entire watershed (Li et al., 2017). Integrating the knowledge on these hot-spots and heterogeneity in general is a question of translating and scaling of specific features and relationships. This has been achieved for carbon stock hot spots combining soil texture and moisture, with PTFs (Bhatti et al., 2002), and also with integration of land use and specific management types (Lal, 2005). Specific investigations to the influence of particle size distribution heterogeneity on soil bulk density values also showed ways of improvement through integration of a Shannon information entropy metric of soil texture in existing PTFs (Martín et al., 2017). This entropy index and soil texture heterogeneity in general might furthermore be estimated based on general lithological character enhancing the possibility to apply such measures more widely (Cámara et al., 2017).

Where it can be considered as a methodological issue, it remains an important challenge to identify sufficiently robust empirical relationships (objective of PTFs) for such hot spots and for processes with hot moments. This points at the need for improved parameterization and validation where a vast array of remote and proximal sensing methods provide now new opportunities (see section 6).

5.4. Global Change Influences on Soil Processes

With atmospheric CO_2 expected to double by 2100, provoking strong changes in hydrological cycles and vegetation cover, CO_2 consumption rates on carbonaceous shale are predicted to increase by more than three times (Davidson & Janssens, 2006). Recent studies further demonstrate that chemical weathering can be highly sensitive to climate changes (Beaulieu et al., 2012). The predicted response to the atmospheric carbon doubling stresses the importance of including chemical weathering in understanding the evolution of global carbon cycle under changing climate conditions (Goddéris & Brantley, 2013). Evidenced relationships between lithology and soil properties should allow for the development of PTFs to predict rock weathering influence on CO_2 consumption. Furthermore, there are PTFs at hand for the interaction between soil-biosphere-atmosphere in diffusion schemes to predict climate-induced changes (Decharme et al., 2011).

It is long time understood that ecosystem models that simulate responses of photosynthesis and respiratory processes to elevated atmospheric CO_2 and increased temperature are fundamental to projecting carbon balance and impacts of global change on the biosphere (Bernacchi et al., 2001; Lloyd & Taylor, 1994; Long, 1991; Sellers et al., 1997). Photosynthetic activity—the terrestrial photosynthetic CO_2 assimilation which can be estimated through soil water content and vapor pressure deficit (Rogers et al., 2016) and thus potentially estimated with PTFs—is directly influenced by the climatic changes and predictions on larger scales (Bernacchi et al., 2001), but up to now without clear accuracy to inform global Earth system models (Rogers et al., 2016). Most terrestrial ecosystem models are able to predict the basic processes of photosynthesis, respiration, and plant growth, yet interactions with the soil nutrient cycle or feed backs from climate change effects are still mostly lacking (Xu-Ri & Prentice, 2008). Temperature sensitivity of soil heterotrophic respiration (i.e., Q10 temperature coefficient measure of the rate of change of a biological or chemical system as a consequence of increasing the temperature by 10°C) is another critical parameter regulating carbon-climate feedback. Although many ecosystem models commonly use a constant Q10 (Chen & Tian, 2005; Schimel et al., 2000; Tian et al., 1999), a small deviation of the variable Q10 will significantly change the estimate of the total CO_2 efflux from soil to the atmosphere (Xu & Qi, 2001). This temperature sensitivity of soil heterotrophic respiration has recently been assessed globally (Zhou et al., 2009).

Also, for basic hydraulic properties and dynamics in LSMs, estimation in the light of climatic changes is still quite uncertain (Fisher et al., 2005). Meteorological forcing to evapotranspiration PTFs at a global scale to predict climate change effects have been documented (Tang et al., 2015), but such large-scale exercises are still rare and uncertainties remain high. Fierer et al. (2009) and Xu et al. (2013) present global analyses of soil microbial responses and relationships to soil organic carbon, nitrogen, and phosphorus in terrestrial ecosystems. Perspectives to explore PTFs are highlighted by recent work quantifying global soil carbon losses in response to warming, based on the newly available Soilgrids information (Crowther et al., 2016). Their analysis might be extended to soil carbon stocks deeper than the upper 10 cm based on existing knowledge and application of PTFs.

The dependence of soil C stabilization and decomposition on soil mineral composition has not been considered in models until recently (Sulman et al., 2014). Carbon decomposition is commonly described with first-order

dependence on soil moisture and/or temperature (Manzoni & Porporato, 2009), possibly with an extra Michaelis-Menten term to account for microbial processes (Wieder et al., 2013). The first-order dependence on soil moisture can reproduce C behavior under steady state moisture conditions but not under dynamic conditions such as pulsed rewetting perturbations (Lawrence et al., 2009). Recent studies have recognized this inadequacy in the representation of complex soil C processes, especially when predicting the soil C losses (Rey et al., 2017) and soil C climate feedbacks (Li et al., 2017).

5.5. Fields to Explore

Both in techniques of acquisition of soil information, and in model elements to be parameterized, there are new fields of opportunity for the development of improved PTFs. Based on the reviewed literature here, we can state that the major opportunities that need to be explored are for the development and use of regionalized, scaled, and globally applicable PTFs in Earth System models in order to account for soil diversity, heterogeneity in space and time and specific environmental conditions (e.g., tropics and freeze-thaw cycles).

5.5.1. Soil Spectroscopy

New sources of PTF inputs stem from technology advances in remote and proximal sensing. Development in soil spectroscopy provides such an interesting source of information to be used within a PTF framework. It allows a rapid acquisition of soil information based on the spectral signature of soil materials, which present a unique reflection depending on their molecular structure. McBratney et al. (2006) explore the idea of using spectral data within a soil inference system, which produces predictions based on PTFs.

Spectral data present a large amount of information, and current research focuses on obtaining soil information from visible, near-, and middle-infrared spectra. Fundamental soil properties, such as texture, C content, CEC, pH, and others, can be estimated with a high degree of accuracy. These spectral estimates of soil properties can be used as inputs to various physical and biochemical PTFs as outlined above.

5.5.2. Novel Integration Domains for PTFs: Global Land Use Prediction, Soil Erosion, and Soil-Landscape Development Models

Land use is often subject of global change predictions, yet mostly with rough estimations and moreover irrelevant projections over territories without accounting for geographical and land suitability information. Models have recently been proposed for land use change based on spatial inferences (Hany & Cohen, 2015) that could be extended with soil information to build PTFs for Earth system modeling purposes. The same goes for landscape evolution models, which run on basic equations that contain soil thickness, texture, and structure (Minasny et al., 2015). Also, soil erosion models are recently explored for approaches to integrate the horizontal (spatial) and vertical processes and first model approaches with PTFs successfully couple erosion and overland flow (Guber et al., 2014). The potential for development of PTFs can be seen in recent global soil erosion model adjustments with soil erodibility estimated based on soil properties of sand, silt and clay fractions, organic matter percent, and gravel percent (Naipal et al., 2015). These relationships between soil properties and soil erodibility are based on recent works exploring the variability across soil types globally (Cerdan et al., 2010; Doetterl et al., 2012).

Applying suites of models with coupled parameterization (Duffy et al., 2014), allow integrating geomorphological, hydrological, and biogeochemical models into LSMs (Brantley et al., 2017). Still, clearly much work is required to better model the effect of climate and organisms on the soil chemistry, mineralogy, physics, and biology (Chadwick et al., 2003). In conclusion, there is still an essential model coupling needed to integrate soil processes, which are usually only represented at a profile scale, with landscape processes (Minasny et al., 2015).

6. Outlook

6.1. Perspectives in Methodological Advances and Dealing With Uncertainty

Since geomorphic information has long been routinely used in soil mapping, geomorphometry may be a valuable data source to predict basic soil properties. In particular, soil texture, organic matter content, and bulk density are known to reflect both landscape position and land surface shape. Because these soil properties are most often included in PTFs, one can hypothesize that soil hydraulic properties should have some relationship to landscape position and land surface shape. Overall, the reported studies and related discussions confirm the usefulness of topographic attributes as ancillary data to indirectly estimate soil properties.

In order to account for soil diversity and spatial heterogeneity in the here crucially identified extrapolation and scaling, the nonlinearity of responses has to be treated in the development of PTFs in the perspective of global model applications. The use of high-resolution soil profile data can be combined with environmental variables (topography, climate, and land cover) through different model approaches for developing broader applicable PTFs to predict the spatial variability of, for example, soil properties or environmental forcing. Difficulty for extrapolation and upscaling lies in the nonlinearity of environmental relationships entered in PTFs.

Therefore, major methodological improvement can be expected first in the replacement of linear models with nonlinear models to account for nonlinear relationships—such as for soil property-depth relationships—and also more complex models to better represent soil-covariate relationships. With an improved understanding of environmental drivers (covariates) for the construction of PTFs and less restrictions of computing capacities, better estimations can be obtained including more covariates in more complex PTFs. Second, improvement can also be expected by the replacement of single prediction models with an ensemble framework, that is, using at least two methods for each soil variable, in order to reduce overshooting effects. Although PTFs in their origin and essence are simple rules, in view of the complexity we can handle nowadays, even a multiple and multivariate model formulation that enables estimating soil properties and processes, responds to the call for PTFs applicable in Earth system science.

Novel (remote) sensing techniques offer a real breakthrough, not only for the development of PTFs but also in the possibilities for model validation. The vast array of remote and proximal sensing methods provides now new opportunities for estimation and validation of relationships between properties, states, and parameters. The importance of these novel sensing techniques, hand in hand with development of PTFs and other spatial inferences, is undeniably at the basis of the huge progress in Earth system modeling. Geophysical techniques, such as ground-penetrating radar, penetrometers, electrical conductivity surveys, and remote geophysical sensing, provide spatial coverage that shows great potential to be incorporated in future PTF developments and applications (Mohanty, 2013). In the near future another growing source of information might be spectral reflectance parameters, which can provide the possibility to gain more information about soils. Therefore, a further possibility to provide more accurate information at catchment and regional scale is to derive prediction methods that rely on spectral data, possibly combined with other environmental data, thereby deriving spectral pedotransfer functions (Babaeian Homaee, Vereecken, et al., 2015).

Measurement uncertainty of soil properties varies (Vereecken, 2002) and can greatly influence model uncertainty. It would be desirable to consider it more explicitly in the development of PTFs to provide more realistic information about the deviation between measured and predicted values. To quantify measurement uncertainty, it is indispensable to define the representative elementary volume of the dependent variable. Information on sample dimension—for example, height or diameter—of the response properties can further improve the prediction accuracy (Ghanbarian et al., 2015), because it provides information on the support of the data from the scale triplet.

The uncertainty in estimated soil properties using PTFs at the model grid scale is composed of three potential sources: (1) the uncertainty in information about texture and other input data, (2) the uncertainty in estimated soil properties propagated through the uncertainty in the PTFs, and (3) the uncertainty induced by aggregating soil information to the grid scale of models. To our knowledge, the impact of uncertainty in PTFs and the different types of PTFs used to predict fluxes and states have not yet been evaluated in global climate models. Also, the influence of the different aggregation/upscaling approaches—Pachepsky and Hill (2017) overviewed concepts of scaling methods in soil science—in global climate models needs clarification. The upscaling of local-scale soil properties used in the various PTFs to predict soil water and energy-related fluxes at the global scale has not been fully clarified especially in view of the newly available high-quality global soil maps.

6.2. Outlook Directions

The identified ways of improving the development and application of PTFs in Earth system science can be represented under three axes: the identification (axis 1), validation (axis 2), and integration (axis 3) of relationships between soil properties, states, and process parameters (Figure 7). The first axis of improvement lies in bridging knowledge gaps in soil process parameterization. Many unknowns, for example, are still present to
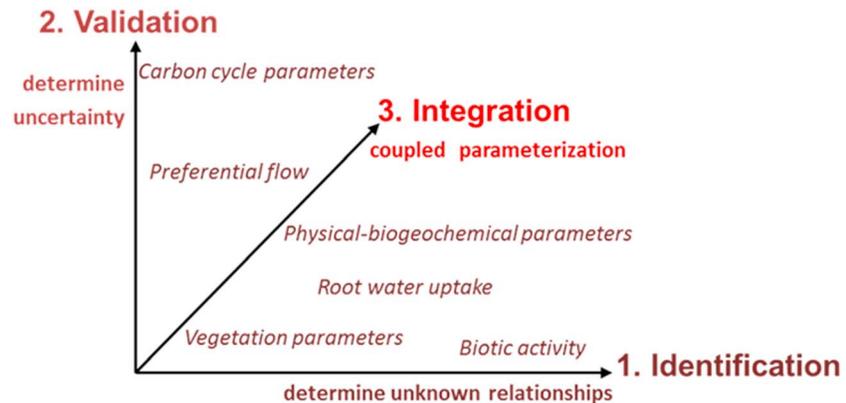
Figure 7. Outlook scheme for PTF development and application of suites of PTFs in Earth system modeling.

relate biotic processes and parameters to soil properties, states, and parameters. These challenges were addressed in section 5 where we identified new pathways for PTFs in Earth system modeling. Especially, the default parameters in models such as soil respiration Q10 or vegetation parameters are in need of validation. The second axis for improvement is present in the validation of identified relationships to a more general application in large-scale models. The methodological challenges (see section 3) for extrapolation and upscaling are described here, and many examples of recent developments in this field are given, such as the estimation of preferential flow and carbon stocks. Finally, for integration especially the linkages between hydraulic and biogeochemical parameters offer strong perspectives currently. Here not only methodological improvements (see section 3.3) but also experimental studies are necessary to derive PTFs for complex models of solute transport including biotic adsorption (see section 4.2) or boundary layer heat exchange processes (see section 4.3).

This scheme can also be presented as a framework for a stepwise development of an ensemble or suite of PTFs for Earth system model applications. The first axis represents the first step in improving the better integration of soil information in Earth system sciences. The second step depicts the necessary step of spatially deploying this knowledge (PTF validation at regional to global scales) and third the integration and linking of the complex model parameterizations to be approached. In this last step, current developments of coupled nitrogen-carbon ESMs show the high potential for the proposed suite of PTFs in the light of predicting gross ecosystem production, respiration, and carbon stocks. Such integration approaches are currently also proposed in global programs to improve carbon stock estimation. Recently, the Global Soil Partnership launched a global endeavor to develop a global soil organic carbon map by the end of 2017, with a soil organic carbon mapping cookbook—with integration recipes (http://www.fao.org/3/a-bs901e.pdf).

Other applications of such suites of PTFs can be for land evaluation and crop productivity or for land protection and sustainability. An overview of soil models using pedotransfer functions with specific application to quantifying ecosystem services is already given in Vereecken et al. (2016). With the here described examples of coupled C-N global model applications and vegetation dynamics parameterizations, specific Earth system research topics can be positioned in this scheme for their prioritization. The process parameterization of preferential flow has priority between the second and third axis, for validation of parameterization and coupling of physical with biogeochemical estimations. A process like root water uptake (or Earth surface exchange processes) needs more input from identifying specific relationships with soil biota in relationship and linking with hydraulic parameters or root zone water storage capacity, driven by rooting depths, needs more model-derived estimations to determine its influence to catchment scale hydrology (Gao et al., 2014).

Integration and validation of existing and novel information is in general a crucial but often retarded process in Earth system modeling. Most of the currently applied parametric PTFs do not yet account for improved soil water flow models (for MRC and HCC), such as presented by Ippisch et al. (2006) or Schaap and van Genuchten (2006), which might be incorporated in land surface models in the future. Further research in prediction of the parameters of bimodal functions (Li et al., 2014) would also yield desirable progress in increasing the performance of Earth system models.

6.3. Time-Dependent PTFs

It is well known that many soil properties vary not only in space but also in time. At present almost all of the PTFs developed up to now assume that predictors remain constant in time. A first approach toward developing time-dependent PTFs (TPTFs) would be to include time-dependent predictor variables (Vereecken et al., 2010). Classical time-dependent predictors in regression-based equations include, for example, bulk density, pH, or vegetation parameters (LAI and NDVI), to estimate, for example, water content and MvG parameter. TPTFs would also be of use for a better acknowledgement of soil structure; for example, in systems with tillage (Kargas et al., 2016), clay soils with swell/shrink behavior (Brake et al., 2013), or soils with proliferic root growth (Fisher et al., 2015). Despite the need for development, TPTFs seem to be challenging. An a priori estimation of the expected temporal variation at scales relevant for Earth system models is rather difficult, as the extent of the variability is unknown and cannot adequately be determined in the laboratory.

In situ estimation of the soil hydraulic parameters by means of data assimilation—combining measurements and models and including all corresponding uncertainties—has been accomplished by Bauser et al. (2016) for a 1-D instrumented field site, with the exception of the local equilibrium assumption during a rain event. Using such an approach for larger-scale assessment is difficult, but some attempts have been done on assimilated soil water content for soil hydraulic parameter estimation (Han et al., 2014; Montzka et al., 2011). Another data-driven approach was presented by Over et al. (2015) using the method of anchored distributions (Rubin et al., 2010), which is a stochastic Bayesian approach using a data-driven and assumption-free likelihood function.

One way for developing TPTFs in Earth systems modeling may be to use the current available posteriori data on changes in PTFs. Climate manipulation sites have been used in long year experiments that may reveal interesting irreversible changes in the soil structure and hence in related PTF values (Robinson et al., 2016). Another way may be to use multiple PTFs to get an ensemble of possible responses, as, for example, shown by Ramírez et al. (2017) for assessing the effect of land use changes in tropical mountain cloud forests. Furthermore, assessing the variability in existing data sets, either by traditional methods or by use of large-scale data sets obtained from satellite remote sensing (Babaeian, Homaee, Montzka, et al., 2015; Paloscia et al., 2013) in combination with soil maps (Levi et al., 2015), may be instrumental in determining a probability density function of expected changes in TPTFs.

Another time-related property at larger scales signifying time dependency of PTFs is the response lag of a catchment or Earth system in question. The challenge herein lies in biophysical feedbacks that are asynchronous, with time lags in system reponse that need to be attributed to specific time-dependent variables in the PTFs.

7. Conclusions

In this contribution we outline the perspectives for the development and application of PTFs with specific focus on improving parameterizations of soil processes in Earth system models. The Earth surface governs ecosystem production, carbon sequestration, and nutrient cycling and controls the heat, moisture, and greenhouse gas exchanges with the atmosphere, which can be determined by the physical properties of the soil and soil state variables through PTFs. The review of current advancements in use of PTFs in Earth system sciences identified the following highlights and targets for further study:

1. A large body of literature exists on PTFs for water, solute, heat, and biogeochemical soil processes applicable for Earth system models, with some recent developments in the fields of root water uptake, solute transport constrained by biogeochemical processes and preferential flows, mineralization, and decomposition processes.
2. Methodological advances are still needed for applications in Earth system models: state-of-the-art extrapolation and upscaling methods are promising, with techniques like geographically weighted regressions to couple PTF-derived parameters with more accurate topographical and geographical information for spatial extrapolation.
3. Challenges for PTF development are determined along three axes: first taking benefit from increased process understanding and determined relationships to enhance parameter estimations. This first axis is furthermore boosted by the newly available high-resolution soil information, and its applications are mainly present in the improved opportunity to incorporate biogeochemical and biotic processes in

Earth system modeling. This enables evaluation of currently set constants (default parameters) that in reality appear to be spatially and temporally variable. Here again, the increased information and high-resolution data allow for validation and to develop improved knowledge rules, as demonstrated for the Q10 relationship and the processes of soil respiration, with implications for climate change prediction in Earth system models. Examples of current developments are the replacement of plant functional types (or completing) with PTFs for some crucial elements like rooting depth and density, to estimate water uptake. It appears feasible and notably better to use models to simulate the root distribution, applying PTFs than to use one value for a plant functional type. This might offer better spatially differentiated estimates of water uptake, caused by the variability of root depths within a plant functional type, and provides better predictions of evapotranspiration. Plant functional types, as commonly used in vegetation models, would better be represented by state variables rather than fixed parameter values (Kattge et al., 2011), and in combination with soil information, more relevant plant functional parameters (for root density and vegetation water content) can be derived.

4. The second axis of PTF development refers to the validation of developed PTFs for large-scale applications. Here both methodological advances of extrapolation and upscaling, as well as newly derived global soil hydraulic functions with their uncertainties, need to be presented and validated.

5. The third axis or step in the framework is to further integrate knowledge in Earth system modeling, with the development of (coupled) suites of PTFs that can be used for solving complex biogeochemical processes in land surface models. Clear perspectives are present for applying suites of models with coupled parameterization, coupling geographical and soil property and state descriptors with the different hydrological and biogeochemical processes in LSMs and also for applications addressing pressing environmental issues like carbon sequestration, soil ecosystem services, and sustainability.

6. Finally, perspectives for novel methods are provided to retrieve detailed soil information (such as spectroscopy and novel remote sensing platforms) applicable in PTF development and model validation and dealing with uncertainty in estimation.

References

Achat, D. L., Pousse, N., Nicolas, M., Bredoire, F., & Augusto, L. (2016). Soil properties controlling inorganic phosphorus availability: General results from a national forest network and a global compilation of the literature. *Biogeochemistry*, *127*(2-3), 255–272. https://doi.org/10.1007/s10533-015-0178-0

Adams, W. A. (1973). The effect of organic matter on the bulk and true densities of some uncultivated podzolic soils. *Journal of Soil Science*, *24*, 10–17.

Adhikari, K., & Hartemink, A. E. (2016). Linking soils to ecosystem services—A global review. *Geoderma*, *262*, 101–111. https://doi.org/10.1016/j.geoderma.2015.08.009

Agyare, W. A., Park, S. J., & Vlek, P. L. G. (2007). Artificial neural network estimation of saturated hydraulic conductivity. *Vadose Zone Journal*, *6*(2), 423–431. https://doi.org/10.2136/vzj2006.0131

Ahuja, L. R., Naney, J. W., Green, R. E., & Nielsen, D. R. (1984). Macroporosity to characterize spatial variability of hydraulic conductivity and effects of land management. *Soil Science Society of America Journal*, *57*, 699–702.

Ahuja, L. R., Naney, J. W., & Williams, R. D. (1985). Estimating soil water characteristics from simpler properties or limited data. *Soil Science Society of America Journal*, *49*(5), 1100–1105. https://doi.org/10.2136/sssaj1985.03615995004900050005x

Akpa, S. I. C., Ugbaje, S. U., Bishop, T. F. A., & Odeh, I. O. A. (2016). Enhancing pedotransfer functions with environmental data for estimating bulk density and effective cation exchange capacity in a data-sparse situation. *Soil Use and Management*, *32*(4), 644–658. https://doi.org/10.1111/sum.12310

Ali, M. H., & Biswas, T. D. (1968). Soil water retention and release as related to mineralogy of soil clays, Proc. 55th Indian Science Congress, 3, 633.

Al-Qinna, M. I., & Jaber, S. M. (2013). Predicting soil bulk density using advanced pedotransfer functions in an arid environment. *Transactions of the ASABE*, *56*(6), 963–976.

Arrouays, D., Grundy, M. G., Hartemink, A. E., Hempel, J. W., Heuvelink, G. B., Young Hongjj, S., … Zhang, G.-l. (2014). Chapter three—GlobalSoilMap: Toward a fine-resolution global grid of soil properties. In D. L. Sparks (Ed.), *Advances in Agronomy* (Vol. 125, pp. 93–134). Soil carbon, Academic Press.

Assouline, S., & Or, D. (2013). Conceptual and parametric representation of soil hydraulic properties: A review. *Vadose Zone Journal*, *12*(4), 1–20. https://doi.org/10.2136/vzj2013.07.0121

Assouline, S., & Or, D. (2014). The concept of field capacity revisited: Defining intrinsic static and dynamic criteria for soil internal drainage dynamics. *Water Resources Research*, *50*(6), 4787–4802. https://doi.org/10.1002/2014wr015475

Babaeian, E., Homaee, M., Vereecken, H., Montzka, C., Norouzi, A. A., & van Genuchten, M. T. (2015). A comparative study of multiple approaches for predicting the soil–water retention curve: Hyperspectral information vs. basic soil properties. *Soil Science Society of America Journal*, *79*(4), 1043–1058. https://doi.org/10.2136/sssaj2014.09.0355

Babaeian, E., Homaee, M., Montzka, C., Vereecken, H., & Norouzi, A. A. (2015). Towards retrieving soil hydraulic properties by hyperspectral remote sensing. *Vadose Zone Journal*, *14*(3). https://doi.org/10.2136/vzj2014.07.0080

Baker, F. G. (1978). Variability of hydraulic conductivity within and between nine Wisconsin soil series. *Water Resources Research*, *14*(1), 103–108. https://doi.org/10.1029/WR014i001p00103

Baker, F. G., & Bouma, J. (1975). Variability of hydraulic conductivity in two subsurface horizons of two siltloam soils. *Soil Science Society of America Journal*, *40*, 219–222.

Baker, L., & Ellison, D. (2008a). Optimisation of pedotransfer functions using an artificial neural network ensemble method. *Geoderma*, *144*(1-2), 212–224. https://doi.org/10.1016/j.geoderma.2007.11.016

Baker, L., & Ellison, D. (2008b). The wisdom of crowds—Ensembles and modules in environmental modelling. *Geoderma*, *147*(1-2), 1–7. https://doi.org/10.1016/j.geoderma.2008.07.003

Bardgett, R. D., & van der Putten, W. H. (2014). Belowground biodiversity and ecosystem functioning. *Nature*, *515*(7528), 505–511. https://doi.org/10.1038/nature13855

Barot, S., Rossi, J. P., & Lavelle, P. (2007). Self-organization in a simple consumer-resource system, the example of earthworms. *Soil Biology and Biochemistry*, *39*(9), 2230–2240. https://doi.org/10.1016/j.soilbio.2007.03.021

Batjes, N. H. (2012). ISRIC-WISE derived soil properties on a 5 by 5 arc-minutes global grid (ver. 1.2). Report 2012/01. Wageningen: ISRIC — World Soil Information (57 pp.).

Batjes, N. H., Ribeiro, E., van Oostrum, A., Leenaars, J., Hengl, T., & Mendes de Jesus, J. (2017). WoSIS: Providing standardised soil profile data for the world. *Earth System Science Data*, *9*(1), 1–14. https://doi.org/10.5194/essd-9-1-2017

Batlle-Aguilar, J., Brovelli, A., Porporato, A., & Barry, D. A. (2011). Modelling soil carbon and nitrogen cycles during land use change. A review. *Agronomy for Sustainable Development*, *31*(2), 251–274. https://doi.org/10.1051/agro/2010007

Bauer, J., Weihermüller, L., Huisman, J. A., Herbst, M., Graf, A., Séquaris, J. M., & Vereecken, H. (2012). Inverse determination of heterotrophic soil respiration response to temperature and water content under field conditions. *Biogeochemistry*, *108*(1–3), 119–134. https://doi.org/10.1007/s10533-011-9583-1

Baumer, O. W. (1992). Predicting unsaturated hyfraulic parameters. In F. J. Leij & L. J. Lund (Eds.), *Indirect methods for estimating the hydraulic properties of unsaturated soils* (pp. 341–355). Riverside, CA: University of California Press.

Bauser, H. H., Jaumann, S., Berg, D., & Roth, K. (2016). EnKF with closed-eye period—Towards a consistent aggregation of information in soil hydrology. *Hydrology and Earth System Sciences*, *20*(12), 4999–5014. https://doi.org/10.5194/hess-20-4999-2016

Bayat, H., Ebrahimi, E., Ersahin, S., Hepper, E. N., Singh, D. N., Amer, A. M., & Yukselen-Aksoy, Y. (2015). Analyzing the effect of various soil properties on the estimation of soil specific surface area by different methods. *Applied Clay Science*, *116-117*, 129–140. https://doi.org/10.1016/j.clay.2015.07.035

Beaulieu, E., Goddéris, Y., Donnadieu, Y., Labat, D., & Roelandt, C. (2012). High sensitivity of the continental-weathering carbon dioxide sink to future climate change. *Nature Climate Change*, *2*(5), 346–349. https://doi.org/10.1038/nclimate1419

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., … Papale, D. (2010). Terrestrial gross carbon dioxide uptake: Global distribution and covariation with climate. *Science*, *329*(5993), 834–838. https://doi.org/10.1126/science.1184984

Behrens, T., & Scholten, T. (2006). Digital soil mapping in Germany—A review. *Journal of Plant Nutrition and Soil Science*, *169*(3), 434–443. https://doi.org/10.1002/jpln.200521962

Bell, M. A., & van Keulen, H. (1995). Soil pedotransfer functions for four Mexican soils. *Soil Science Society of America Journal*, *59*(3), 865–871. https://doi.org/10.2136/sssaj1995.03615995005900030034x

Bender, S. F., Wagg, C., & van der Heijden, M. G. A. (2016). An underground revolution: Biodiversity and soil ecological engineering for agricultural sustainability. *Trends in Ecology & Evolution*, *31*(6), 440–452. https://doi.org/10.1016/j.tree.2016.02.016

Bengough, A. G., Bransby, M. F., Hans, J., McKenna, S. J., Roberts, T. J., & Valentine, T. A. (2006). Root responses to soil physical conditions; growth dynamics from field to cell. *Journal of Experimental Botany*, *57*(2), 437–447. https://doi.org/10.1093/jxb/erj003

Benites, V. M., Machado, P. L. O. A., Fidalgo, E. C. C., Coelho, M. R., & Madari, B. E. (2007). Pedotransfer functions for estimating soil bulk density from existing soil survey reports in Brazil. *Geoderma*, *139*(1-2), 90–97. https://doi.org/10.1016/j.geoderma.2007.01.005

Beringer, J., Lynch, A. H., Chapin II, F. S., Mack, M., & Bonan, G. B. (2001). The representation of Arctic soils in the land surface model: The importance of mosses. *J Climate*, *14*, 3324–3335.

Bernacchi, C. J., Singsaas, E. L., Pimentel, C., Portis, A. R. Jr., & Long, S. P. (2001). Improved temperature response functions for models of Rubisco-limited photosynthesis. *Plant, Cell & Environment*, *24*(2), 253–259. https://doi.org/10.1111/j.1365-3040.2001.00668.x

Bernoux, M., Arrouays, D., Cerri, C., Volkoff, B., & Jolivet, C. (1998). Bulk densities of Brazilian Amazon soils related to other soil properties. *Soil Science Society of America Journal*, *62*(3), 743–749. https://doi.org/10.2136/sssaj1998.03615995006200030029x

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Menard, C. B., … Harding, R. J. (2011). The Joint UK Land Environment Simulator (JULES), model description—Part 1: Energy and water fluxes. *Geoscientific Model Development*, *4*(3), 677–699. https://doi.org/10.5194/gmd-4-677-2011

Bevington, J., Piragnolo, D., Teatini, P., Vellidis, G., & Morari, F. (2016). On the spatial variability of soil hydraulic properties in a Holocene coastal farmland. *Geoderma*, *262*, 294–305. https://doi.org/10.1016/j.geoderma.2015.08.025

Bhatti, J. S., Apps, M. J., Tarnocai, C., Karjalainen, T., Stocks, B. J., & Shaw, C. (2002). Estimates of soil organic carbon stocks in central Canada using three different approaches. *Canadian Journal of Forest Research*, *32*(5), 805–812. https://doi.org/10.1139/x01-122

Billen, N., Röder, C., Gaiser, T., & Stahr, K. (2009). Carbon sequestration in soils of SW-Germany as affected by agricultural management-calibration of the EPIC model for regional simulations. *Ecological Modelling*, *220*(1), 71–80. https://doi.org/10.1016/j.ecolmodel.2008.08.015

Blanchart, E., Marilleau, N., Chotte, J. L., Drogoul, A., Perrier, E., & Cambier, C. (2009). SWORM: An agent-based model to simulate the effect of earthworms on soil structure. *European Journal of Soil Science*, *60*(1), 13–21. https://doi.org/10.1111/j.1365-2389.2008.01091.x

Bloemen, G. W. (1977). Calculation of capillary conductivity and capillary rise from grainsize distribution. ICWWageningen note no. 952, 962, 1013.

Bloemen, G. W. (1980). Calculation of hydraulic conductivities from texture and organic matter content. *Zeitschrift für Pflanzenernährung und Bodenkunde*, *143*(5), 581–605. https://doi.org/10.1002/jpln.19801430513

Bonan, G. B., & Levis, S. (2010). Quantifying carbon-nitrogen feedbacks in the Community Land Model. (CLM4). *Geophysical Research Letters*, *37*, L07401. https://doi.org/10.1029/2010GL042430

Bormann, H., & Klaasen, K. (2008). Seasonal and land use dependent variability of soil hydraulic and soil hydrological properties of two northern German soils. *Geoderma*, *145*(3-4), 295–302. https://doi.org/10.1016/j.geoderma.2008.03.017

Boschi, R. S., Rodrigues, L. H. A., & Lopes-Assad, M. L. R. C. (2014). Using classification trees to evaluate the performance of pedotransfer functions. *Vadose Zone Journal*, *13*(8). https://doi.org/10.2136/vzj2013.11.0195

Boschi, R. S., Rodrigues, L. H. A., & Lopes-Assad, M. L. R. C. (2015). Analysis of patterns of pedotransfer function estimates: An approach based on classification trees. *Soil Science Society of America Journal*, *79*(3), 720–729. https://doi.org/10.2136/sssaj2014.11.0452

Botula, Y.-D., Nemes, A., Mafuka, P., Van Ranst, E., & Cornelis, W. M. (2013). Prediction of water retention of soils from the humid tropics by the nonparametric-nearest neighbor approach. *Vadose Zone Journal*, *12*(2). https://doi.org/10.2136/vzj2012.0123

Botula, Y. D., Nemes, A., Ranst, E. V., Mafuka, P., Pue, J. D., & Cornelis, W. (2015). Hierarchical pedotransfer functions to predict bulk density of highly weathered soils in central Africa. *Soil Science Society of America Journal*, *79*(2), 476–486. https://doi.org/10.2136/sssaj2014.06.0238

Bouma, J. (1988). When the mapping is over, then what? In: Minn. Extension service, Univ. Minnesota. Proc. of Int. Interactive workshop on soil resources: Their inventory, analysis and interpretation for use in the 1990's, 3–11.

Bouma, J. (1989). Using soil survey data for quantitative land evaluation. In B. A. Stewart (Ed.), *Advances in soil science* (Vol. 9, pp. 177–213). New York: Springer Verlag. https://doi.org/10.1007/978-1-4612-3532-3_4

Bouma, J., Stoorvogel, J. J., & Sonneveld, W. M. P. (2012). Land evaluation for landscape units. In P. M. Huang, Y. Li, & M. Summer (Eds.), *Handbook of soil science* (2nd ed., Vol. 34, pp. 1–22). Boca Raton, London, New York: CRC Press.

Bradford, M. A., Wood, S. A., Bardgett, R. D., Black, H. I. J., Bonkowski, M., Eggers, T., … Jones, T. H. (2014). Discontinuity in the responses of ecosystem processes and multifunctionality to altered soil community composition. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(40), 14,478–14,483. https://doi.org/10.1073/pnas.1413707111

Brake, B. T., van der Ploeg, M. J., & de Rooij, G. H. (2013). Water storage change estimation from in situ shrinkage measurements of clay soils. *Hydrology and Earth System Sciences*, *17*(5), 1933–1949. https://doi.org/10.5194/hess-17-1933-2013

Brantley, S. L., Lebedeva, M. I., Balashov, V. N., Singha, K., Sullivan, P. L., & Stinchcomb, G. (2017). Toward a conceptual model relating chemical reaction fronts to water flow paths in hills. *Geomorphology*, *277*, 100–117. https://doi.org/10.1016/j.geomorph.2016.09.027

Breeuwsma, A., Wösten, J. H. M., Vleeshouwer, J. J., van Slobbe, A. M., & Bouma, J. (1986). Derivation of land qualities to assess environmental problems from soil surveys. *Soil Science Society of America Journal*, *50*(1), 186–190. https://doi.org/10.2136/sssaj1986.03615995005000010035x

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. Boca Raton, FL: Chapman and Hall/CRC.

Bricklemyer, R. S., Miller, P. R., Turk, P. J., Paustian, K., Keck, T., & Nielsen, G. A. (2007). Sensitivity of the century model to scale-related soil texture variability. *Soil Science Society of America Journal*, *71*(3), 784–792. https://doi.org/10.2136/sssaj2006.0168

Briggs, L. J., & Lane, J. W. M. (1907). The moisture equivalents of soils. (U.S. Department of Agriculture Bureau of Soils. Bulletin) (pp. 23). U.S. Government Printing Office. Retrieved from https://books.google.com/books?id=oWp2rgEACAAJ

Brimelow, J. C., Hanesiak, J. M., & Raddatz, R. (2010). Validation of soil moisture simulations from the PAMII model, and an assessment of their sensitivity to uncertainties in soil hydraulic parameters. *Agricultural and Forest Meteorology*, *150*, 100–114.

Brooks, R. H., & Corey, A. T. (1964). Hydraulic properties of porous media. *Hydrol.*, Paper No. *3*, Colorado State Univ., Fort Collins, CO.

Brusseau, M. L., & Rao, P. S. C. (1990). Modeling the transport of solutes influenced by multiprocess nonequilibrium. *Water Resources Research*, *26*(1), 165. https://doi.org/10.1029/WR026i001p00165

Brutsaert, W. (1966). Probability laws for pore-size distributions. *Soil Science*, *101*(2), 85–92. https://doi.org/10.1097/00010694-196602000-00002

Burdine, N. T. (1953). Relative permeability calculations from pore size distribution06.011, 2010.

Cagliari, J., Veronez, M. R., & Alves, M. E. (2011). Remaining phosphorus estimated by pedotransfer functions. *Revista Brasileira de Ciência do Solo*, *35*(1), 203–212. https://doi.org/10.1590/S0100-06832011000100019

Cale, W. G., Oneill, R. V., & Gardner, R. H. (1983). Aggregation error in non-linear ecological models. *Journal of Theoretical Biology*, *100*(3), 539–550. https://doi.org/10.1016/0022-5193(83)90446-0

Calvet, J. C., Fritz, N., Berne, C., Piguet, B., Maurel, W., & Meurey, C. (2016). Deriving pedotransfer functions for soil quartz fraction in southern France from reverse modeling. *The Soil*, *2*(4), 615–629. https://doi.org/10.5194/soil-2-615-2016

Cámara, J., Gómez-Miguel, V., & Martín, M. Á. (2017). Lithologic control on soil texture heterogeneity. *Geoderma*, *287*, 157–163. https://doi.org/10.1016/j.geoderma.2016.09.006

Camargo, L. A., Marques, J. Jr., Pereirea, G. T., Alleoni, L. R. F., De Bahia, A. S. R., De, D., & Texeira, B. (2016). Pedotransfer functions to assess adsorbed phosphate using iron oxide content and magnetic susceptibility in an Oxisol. *Soil Use and Management*, *32*(2), 172–182. https://doi.org/10.1111/sum.12255

Campbell, G. S. (1974). A simple method for determining unsaturated conductivity from moisture retention data. *Soil Science*, *117*(6), 311–314.

Campbell, G. S., & Shiozawa, S. (1992). Prediction of hydraulic properties of soils using particle-size distribution and bulk density data, Indirect methods Estim. Hydraul. Prop. unsaturated soils (pp. 317–328). University California, Riverside.

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

Carvalhais, N., Reichstein, M., Seixas, J., James Collatz, G., Pereira, J. S., Berbigier, P., … Valentini, R. (2008). Implications of the carbon cycle steady state assumption for biogeochemical modeling performance and inverse parameter retrieval. *Global Biogeochemical Cycles*, *22*, GB2007. https://doi.org/10.1029/2007GB003033

Casanova, M., Tapia, E., Seguel, O., & Salazar, O. (2016). Direct measurement and prediction of bulk density on alluvial soils of central Chile. *Chilean Journal of Agricultural Research*, *76*(1), 105–113. https://doi.org/10.4067/S0718-58392016000100015

Cerdan, O., Govers, G., Le Bissonnais, Y., van Oost, K., Poesen, J., Saby, N., … Dostal, T. (2010). Rates and spatial variations of soil erosion in Europe: A study based on erosion plot data. *Geomorphology*, *122*(1-2), 167–177. https://doi.org/10.1016/j.geomorph

Chadwick, O. A., Gavenda, R. T., Kelly, E. F., Ziegler, K., Olson, C. G., Elliott, W. C., & Hendricks, D. M. (2003). The impact of climate on the biogeochemical functioning of volcanic soils. *Chemical Geology*, *202*, 195–223. https://doi.org/10.1016/j.chemgeo.2002.09.001

Chaney, N. W., Wood, E. F., McBratney, A. B., Hempel, J. W., Nauman, T. W., Brungard, C. W., & Odgers, N. P. (2016). POLARIS: A 30-meter probabilistic soil series map of the contiguous United States. *Geoderma*, *274*, 54–67.

Chen, H., & Tian, H. Q. (2005). Does a general temperature-dependent Q10 model of soil respiration exist at biome and global scale? *Journal of Integrative Plant Biology*, *47*(11), 1288–1302. https://doi.org/10.1111/j.1744-7909.2005.00211.x

Clapp, R. B., & Hornberger, G. M. (1978). Empirical equations for some soil hydraulic properties. *Water Resources Research*, *14*(4), 601–604. https://doi.org/10.1029/WR014i004p00601

Coats, K. H., & Smith, B. D. (1964). Dead-end pore volume and dispersion in porous media. *Society of Petroleum Engineers Journal*, *4*, 73–84.

Cohen, M. J., Paris, J., & Clark, M. W. (2007). P-sorption capacity estimation in southeastern USA wetland soils using visible/near infrared (VNIR) reflectance spectroscopy. *Wetlands*, *27*(4), 1098–1111. https://doi.org/10.1672/0277-5212(2007)27%5B1098:PCEISU%5D2.0.CO;2

Coleman, K. and Jenkinson, D. S. (2005). RothC-26.3. A model for turnover of carbon in soil, model description and windows users guide. IACR-Rothamsted, Harpenden (45 pp). Retrieved from http://www.rothamsted.bbsrc.ac.uk/aen/carbon/rothc.htm

Coleman, K., Jenkinson, D. S., Crocker, G. J., Grace, P. R., Klír, J., Körschens, M., … Richter, D. D. (1997). Simulating trends in soil organic carbon in long-term experiments using RothC-26.3. *Geoderma*, *81*(1–2), 29–44. https://doi.org/10.1016/S0016-7061(97)00079-7

Cornelis, W. M., Ronsyn, J., van Meirvenne, M., & Hartmann, R. (2001). Evaluation of pedotransfer functions for predicting the soil moisture retention curve. *Soil Science Society of America Journal*, *65*(3), 638–648. https://doi.org/10.2136/sssaj2001.653638x

Cosby, B. J., Hornberger, G. M., Clapp, R. B., & Ginn, T. R. (1984). A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resources Research*, *20*, 682–690.

Côté, J., & Konrad, J.-M. (2005). A generalized thermal conductivity model for soils and construction materials. *Canadian Geotechnical Journal*, *42*(2), 443–458. https://doi.org/10.1139/T04-106

Cox, P. M., Betts, R. A., Bunton, C. B., Essery, R. L. H., Rowntree, P. R., & Smith, J. (1999). The impact of new land surface physics on the GCM simulation of climate and climate sensitivity. *Climate Dynamics*, *15*, 183–203.

Crowther, T. W., Todd-Brown, K. E. O., Rowe, C. W., Wieder, W. R., Carey, J. C., Machmuller, M. B., … Bradford, M. A. (2016). Quantifying global soil carbon losses in response to warming. *Nature*, *540*, 104–108.

Dai, Y., Shangguan, W., Duan, Q., Liu, B., Fu, S., & Niu, G.-Y. (2013). Development of a China dataset of soil hydraulic parameters using Pedotransfer functions for land surface modeling. *Journal of Hydrometeorology*, *14*(3), 869–887. https://doi.org/10.1175/JHM-D-12-0149.1

Darrah, P. R., Jones, D. L., Kirk, G. J. D., & Roose, T. (2006). Modelling the rhizosphere: A review of methods for 'upscaling' to the whole-plant scale. *European Journal of Soil Science*, *57*, 13–25. https://doi.org/10.1111/j.1365-2389.2006.00786.x

Davidson, E. A., & Janssens, I. A. (2006). Temperature sensitivity of soil carbon decomposition and feed backs to climate change. *Nature*, *440*(7081), 165–173. https://doi.org/10.1038/nature04514

De Lannoy, G. J. M., Koster, R. D., Reichle, R. H., Mahanama, S. P. P., & Liu, Q. (2014). An updated treatment of soil texture and associated hydraulic properties in a global land modeling system. *Journal of Advances in Modeling Earth Systems*, *6*(4), 957–979. https://doi.org/10.1002/2014MS000330

de Rooij, G. H., Cirpka, O. A., Stagnitti, F., Vuurens, S. H., & Boll, J. (2006). Quantifying minimum monolith size and solute dilution from multi-compartment percolation sampler data. *Vadose Zone Journal*, *5*(4), 1086–1092. https://doi.org/10.2136/vzj2005.0101

De Souza, E., Filho, E. I. F., Schaefer, C. E. G. R., Batjes, N. H., dos Santos, G. R., & Machado Pontes, L. (2016). Pedotransfer functions to estimate bulk density from soil properties and environmental covariates: Rio Doce basin. *Science in Agriculture*, *73*, 525–534.

de Vries, D. A. (1963). Thermal properties of soils. In W. R. Van Wijk (Ed.), *Physics of plant environment* (pp. 210–235). Amsterdam: North-Holland Publ. Co.

de Vries, F. T., Thébault, E., Liiri, M., Birkhofer, K., Tsiafouli, M. A., Bjørnlund, L., … Bardgett, R. D. (2013). Soil food web properties explain ecosystem services across European land use systems. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 14,296–14,301.

Decharme, B., Boone, A., Delire, C., & Noilhan, J. (2011). Local evaluation of the interaction between soil biosphere atmosphere soil multilayer diffusion scheme using four pedotransfer functions. *Journal of Geophysical Research*, *116*, D20126. https://doi.org/10.1029/2011JD016002

Decharme, B., Brun, E., Boone, A., Delire, C., Le Moigne, P., & Morin, S. (2016). Impacts of snow and organic soils parameterization on northern Eurasian soil temperature profiles simulated by the ISBA land surface model. *The Cryosphere*, *10*(2), 853–877. https://doi.org/10.5194/tc-10-853-2016

Dettmann, U., & Bechtold, M. (2016). Deriving effective soil water retention characteristics from shallow water table fluctuations in Peatlands. *Vadose Zone Journal*, *15*(10), 1–13. https://doi.org/10.2136/vzj2016.04.0029

Dexter, A. R., Richard, G., Arrouays, D., Czyż, E. A., Jolivet, C., & Duval, O. (2008). Complexed organic matter controls soil physical properties. *Geoderma*, *144*(3-4), 620–627. https://doi.org/10.1016/j.geoderma.2008.01.022

Dharssi, I., Vidale, P. L., Verhoef, A., Macpherson, B., Jones, C., & Best, M. (2009). New soil physical properties implemented in the Unified Model at PS18 (Tech. Rep. 528), Forecasting R&D, Met Office, Exeter, UK.

Dickinson, R. E., Henderson-Sellers, A., Kennedy, P. J., and Wilson, M. F. (1986). Biosphere–atmosphere transfer scheme (BATS) for the NCAR community climate model (NCAR Tech. Note Tn-275+ STR, 72 pp.).

Dickinson, R. E., Kennedy, P. J., & Henderson-Sellers, A. (1993). Biosphere-atmosphere transfer scheme (BATS) version 1e as coupled to the NCAR community climate model. National Center for Atmospheric Research, Climate and Global Dynamics Division.

Dirmeyer, P., Gao, X., & Oki, T. (2002). GSWP-2: The Second Global Soil Wetness Project Science and Implementation Plan. IGPO Publication Series No. 37, International GEWEX (Global Energy and Water Cycle Experiment) Project Office.

Disse, M. (1995). Modellierung der Verdunstung und der Grundwasserneubildung in ebenen Einzugsgebieten. [Modeling groundwater recharge in small catchments]. *Mitteilungen des Inst.f. Hydrologie u. Wasserwirtschaft d. Universität Karlsruhe 53*, 95–107.

Doetterl, S., van Oost, K., & Six, J. (2012). Towards constraining the magnitude of global agricultural sediment and soil organic carbon fluxes. *Earth Surface Processes and Landforms*, *37*(6), 642–655. https://doi.org/10.1002/esp.3198

Donatelli, M., Wösten, J. H. M., & Belocchi, G. (2004). Methods to evaluate pedotransfer functions. In Y. A. Pachepsky & W. J. Rawls (Eds.), *Development of pedotransfer functions in soil hydrology* (Vol. 30, pp. 357–411). Amsterdam: Elsevier. https://doi.org/10.1016/S0166-2481(04)30020-6

Dong, Y., McCartneyand, J. S., & Lu, N. (2015). Critical review of thermal conductivity models for unsaturated soils. *Geotechnical and Geological Engineering*, *33*(2), 207–221. https://doi.org/10.1007/s10706-015-9843-2

Ducharne, A. (2016). The hydrol module of ORCHIDEE: Scientific documentation. Retrieved from http://forge.ipsl.jussieu.fr/orchidee/attachment/wiki/Documentation/UserGuide/eqs_hydrol.pdf

Ducharne, A., Koster, R. D., Suarez, M. J., Stieglitz, M., & Kumar, P. (2000). A catchment-based approach to modeling land surface processes in a general circulation model 2. Parameter estimation and model demonstration. *Journal of Geophysical Research*, *105*(D20), 24,838–24,838.

Duffy, C., Shi, Y., Davis, K., Slingerland, R., Li, L., Sullivan, P. L., … Brantley, S. L. (2014). Designing a suite of models to explore critical zone function. *Procedia Earth and Planetary Science*, *10*, 7–15. https://doi.org/10.1016/j.proeps.2014.08.003

Dunbabin, V. M., Postma, J. A., Schnepf, A., Pagès, L., Javaux, M., Wu, L., … Diggle, A. J. (2013). Modelling root-soil interactions using three-dimensional models of root growth, architecture and function. *Plant and Soil*, *372*, 93–124.

Dunne, K. A., & Willmott, C. J. (1996). Global distribution of plant-extractable water capacity of soil. *International Journal of Climatology*, *16*(8), 841–859.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Boca Raton, FL: Chapman and Hall/CRC. https://doi.org/10.1007/978-1-4899-4541-9

Eisenhauer, N., Reich, P. B., & Isbell, F. (2012). Decomposer diversity and identity influence plant diversity effects on ecosystem functioning. *Ecology*, *93*(10), 2227–2240. https://doi.org/10.1890/11-2266.1

Engelbrecht, B. M. J., Comita, L. S., Condit, R., Kursar, T. A., Tyree, M. T., Turner, B. L., & Hubbell, S. P. (2007). Drought sensitivity shapes species distribution patterns in tropical forests. *Nature*, *447*(7140), 80–82. https://doi.org/10.1038/nature05747

Esser, G., Kattge, J., & Sakalli, A. (2011). Feedback of carbon and nitrogen cycles enhances carbon sequestration in the terrestrial biosphere. *Global Change Biology*, *17*(2), 819–842. https://doi.org/10.1111/j.1365-2486.2010.02261.x

Falloon, P., Smith, P., Coleman, K., & Marshall, S. (1998). Estimating the size of the inert organic matter pool from total soil organic carbon content for use in the Rothamsted carbon model. *Soil Biology and Biochemistry*, *30*(8-9), 1207–1211. https://doi.org/10.1016/S0038-0717(97)00256-3

Food and Agriculture Organization (FAO) (2012). *Harmonized world soil database (version 1.2)*. Rome: FAO.

Farkas, C., Gyuricza, C., & Birkás, M. (2006). Seasonal changes of hydraulic properties of a Chromic Luvisol under different soil management. *Biologia*, *61*(Suppl. 19), S344–S348.

Farouki, O. T. (1986). *Thermal properties of soils, series on rock and soil mechanics*, Trans. Tech. Pub, (Vol. 11, p. 136). MA, USA: Rockport.

Fenner, K., Lanz, V. A., Scheringer, M., & Borsuk, M. E. (2007). Relating atrazine degradation rate in soil to environmental conditions: Implications for global fate modeling. *Environmental Science and Technology*, *41*(8), 2840–2846. https://doi.org/10.1021/es061923i

Feoli, E., Pérez-Gómez, R., Oyonarte, C., & Ibáñez, J. J. (2017). Using spatial data mining to analyze area-diversity patterns among soil, vegetation, and climate: A case study from Almería, Spain. *Geoderma*, *287*, 164–169. https://doi.org/10.1016/j.geoderma.2016.09.011

Fierer, N., & Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(3), 626–631. https://doi.org/10.1073/pnas.0507535103

Fierer, N., Strickland, M. S., Liptzin, D., Bradford, M. A., & Cleveland, C. C. (2009). Global patterns in belowground communities. *Ecology Letters*, *12*(11), 1238–1249. https://doi.org/10.1111/j.1461-0248.2009.01360.x

Filser, J., Faber, J. H., Tiunov, A. V., Brussaard, L., Frouz, J., De Deyn, G., … José Jiménez, J. (2016). Soil fauna: Key to new carbon models. *The Soil*, *2*(4), 565–582. https://doi.org/10.5194/soil-2-565-2016

Fischer, C., Tischer, J., Roscher, C., Eisenhauer, N., Ravenek, J., Gleixner, G., & Scheu, S. (2015). Plant species diversity affects infiltration capacity in an experimental grassland through changes in soil properties. *Plant and Soil*, *397*(1-2), 1–16. https://doi.org/10.1007/s11104-014-2373-5

Fisher, J. B., DeBiase, T. A., Qi, Y., Xu, M., & Goldstein, A. H. (2005). Evapotranspiration models compared on a Sierra Nevada forest ecosystem. *Environmental Modelling & Software*, *20*(6), 783–796. https://doi.org/10.1016/j.envsoft.2004.09.009

Fodor, N., Sándor, R., Orfanus, T., Lichner, L., & Rajkai, K. (2011). Evaluation method dependency of measured saturated hydraulic conductivity. *Geoderma*, *165*(1), 60–68. https://doi.org/10.1016/j.geoderma.2011.07.004

Fóti, S., Balogh, J., Herbst, M., Papp, M., Koncz, P., Bartha, S., & Nagy, Z. (2016). Meta-analysis of field scale spatial variability of grassland soil CO_2 efflux: Interaction of biotic and abiotic drivers. *Catena*, *143*, 78–89. https://doi.org/10.1016/j.catena.2016.03.034

Freckman, D. W. (1986). The ecology of dehydration in soil organisms. In A. C. Leopold (Ed.), *Membranes, Metabolism, and Dry Organisms* (pp. 157–168). London: Comstock Publishing Associates.

Fredlund, D. G., & Xing, A. (1994). Equations for the soil-water characteristic curve. *Canadian Geotechnical Journal*, *31*(6), 1026–1026. https://doi.org/10.1139/t94-120

Friedlingstein, P., Meinshausen, M., Arora, V., Jones, C. D., Anav, A., Liddicoat, S., & Knutti, R. (2014). Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *Journal of Climate*, *27*(2), 511–526. https://doi.org/10.1175/JCLI-D-12-00579.1

Gao, H., Hrachowitz, M., Schymanski, S. J., Fenicia, F., Sriwongsitanon, N., & Savenije, H. H. G. (2014). Climate controls how ecosystems size the root zone storage capacity at catchment scale. *Geophysical Research Letters*, *41*(22), 7916–7923. https://doi.org/10.1002/2014GL061668

Gardi, C., Menta, C., Montanarella, L., & Cenci, R. (2008). Main threats to soil biodiversity: the case of agricultural activities impacts on soil microarthropods. In G. Toth, L. Montanarella, & E. Rusco (Eds.), *Threats to soil in Europe* (pp. 100–110). Luxembourg: Office for the Official Publications of the European Communities.

Gardi, C., Jeffery, S., & Saltelli, A. (2013). An estimate of potential threats levels to soil biodiversity in EU. *Global Change Biology*, *19*, 1538–1548. https://doi.org/10.1111/gcb.12159

Gardner, W. R. (1958). Some steady state solutions of the unsaturated moisture flow equation with application to evaporation from a water table. *Soil Science*, *85*(4), 228–232. https://doi.org/10.1097/00010694-195804000-00006

Gelhar, L. W. (1986). Stochastic subsurface hydrology from theory to applications. *Water Resources Research*, *22*(9S), 135S–145S. https://doi.org/10.1029/WR022i09Sp0135S

Gerke, H. H., & Van Genuchten, M. T. (1993). A dual-porosity model for simulating the preferential movement of water and solutes in structured porous media. *Water Resources Research*, *29*(2), 305–319. https://doi.org/10.1029/92WR02339

Ghafoor, A., Jarvis, N. J., Thierfelder, T., & Stenstrom, J. (2011). Measurements and modeling of pesticide persistence in soil at the catchment scale. *Science of the Total Environment*, *409*(10), 1900–1908. https://doi.org/10.1016/j.scitotenv.2011.01.049

Ghanbarian, B., Taslimitehrani, V., Dong, G., & Pachepsky, Y. A. (2015). Sample dimensions effect on prediction of soil water retention curve and saturated hydraulic conductivity. *Journal of Hydrology*, *528*, 127–137. https://doi.org/10.1016/j.jhydrol.2015.06.024

Ghehi, N., Nemes, A., Verdoodt, A., Van Ranst, E., Cornelis, W. M., & Boeckx, P. (2012). Use of the nonparametric nearest neighbor and boosted regression tree techniques to estimate soil bulk density in tropical rainforest soils. *Soil Science Society of America Journal*, *76*(4), 1172–1183. https://doi.org/10.2136/sssaj2011.0330

Gijsman, A. J., Jagtap, S. S., & Jones, J. W. (2002). Wading through a swamp of complete confusion: How to choose a method for estimating soil water retention parameters for crop models. *European Journal of Agronomy*, *18*(1-2), 77–106. https://doi.org/10.1016/S1161-0301(02)00098-9

Glendining, M. J., Dailey, A. G., Powlson, D. S., Richter, G. M., Catt, J. A., & Whitmore, A. P. (2011). Pedotransfer functions for estimating total soil nitrogen up to the global scale. *European Journal of Soil Science*, *62*(1), 13–22. https://doi.org/10.1111/j.1365-2389.2010.01336.x

Goddéris, Y., & Brantley, S. L. (2013). Earthcasting the future critical zone. *Elementa*, *1*, 19. https://doi.org/10.12952/journal.elementa.000019

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*(4), 563–574. https://doi.org/10.1046/j.1365-2389.2001.00409.x

Guber, A. K., Pachepsky, Y. A., van Genuchten, M. T., Simunek, J., Jacques, D., Nemes, A., … Cady, R. E. (2009). Multimodel simulation of water flow in a field soil using pedotransfer functions. *Vadose Zone Journal*, *8*(1), 1–10.

Guber, A. K., Pachepsky, Y. A., Yakirevich, A. M., Shelton, D. R., Whelan, G., Goodrich, D. C., & Unkrich, C. L. (2014). Modeling runoff and microbial overland transport with TanDEM-X DEM model: Accuracy and uncertainty as affected by source of infiltration parameters. *Journal of Hydrology*, *519*, 644–655. https://doi.org/10.1016/j.jhydrol.2014.08.005

Gupta, S., & Larson, W. E. (1979). Estimating soil water retention characteristics from particle size distribution, organic matter percent, and bulk density. *Water Resources Research*, *15*(6), 1633–1635. https://doi.org/10.1029/WR015i006p01633

Guswa, A. J. (2008). The influence of climate on root depth: A carbon cost-benefit analysis. *Water Resources Research*, *44*(2), W02427. https://doi.org/10.1029/2007WR006384

Gutmann, E. D., & Small, E. E. (2007). A comparison of land surface model soil hydraulic properties estimated by inverse modeling and pedotransfer functions. *Water Resources Research*, *43*, W05418.

Hagemann, S. (2002). *An improved land surface parameter dataset for global and regional climate models*. Max-Planck-Institute for Meteorology, Report 336, Hamburg, Germany. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.471.2882&rep=rep1&type=pdf

Hagemann, S., & Dümenil Gates, L. (2003). Improving a subgrid runoff parameterization scheme for climate models by the use of high resolution data derived from satellite observations. *Climate Dynamics*, *21*(3-4), 349–359. https://doi.org/10.1007/s00382-003-0349-x

Haghverdi, A., Cornelis, W. M., & Ghahraman, B. (2012). A pseudo-continuous neural network approach for developing water retention pedotransfer functions with limited data. *Journal of Hydrology*, *442*, 46–54.

Haghverdi, A., Öztürk, H. S., & Cornelis, W. M. (2014). Revisiting the pseudo continuous pedotransfer function concept: Impact of data quality and data mining method. *Geoderma*, *226*, 31–38.

Hallema, D. W., Périard, Y., Lafond, J. A., Gumiere, S. J., & Caron, J. (2015). Characterization of water retention curves for a series of cultivated Histosols. *Vadose Zone Journal*, *14*(6), 0. https://doi.org/10.2136/vzj2014.10.0148

Han, X., Hendricks-Franssen, H. J., Montzka, C., & Verrecken, H. (2014). Soil moisture and soil properties estimation in the community land model with synthetic brightness temperature observations. *Water Resources Research*, *50*(7), 6081–6105. https://doi.org/10.1002/2013WR014586

Hany, N., & Cohen, S. (2015). Predicting 21th century global agricultural land use with a spatially and temporally explicit regression-based model. *Applied Geography*, *62*, 366–376. https://doi.org/10.1016/j.apgeog.2015.05.010

Hare, J. D., Elle, E. E., & van Dam, N. M. (2003). Costs of glandular trichomes in Datura Wrightii: A three year study. *Evolution*, *57*(4), 793–805. https://doi.org/10.1111/j.0014-3820.2003.tb00291.x

Hartnett, D. C., Wilson, G. W., Ott, J. P., & Setshogo, M. (2013). Variation in root system traits among African semi-arid savanna grasses: Implications for drought tolerance. *Austral Ecology*, *38*(4), 383–392. https://doi.org/10.1111/j.1442-9993.2012.02422.x

Hassink, J. (1997). The capacity of soils to preserve organic C and N by their association with clay and silt particles. *Plant and Soil*, *191*(1), 77–87. https://doi.org/10.1023/A:1004213929699

Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The elements of statistical learning data mining, inference, and prediction*. New York.

Haykin, S. (1994). *Neural networks: A comprehensive approach*. New York: IEEE Computer Society Press.

He, J.-Z., Ge, Y., Xu, Z., & Chen, C. (2009). Linking soil bacterial diversity to ecosystem multifunctionality using backward-elimination boosted trees analysis. *Journal of Soils and Sediments*, *9*(6), 547–554. https://doi.org/10.1007/s11368-009-0120-y

Hecht-Nielsen, R. (1990). On the algebraic structure of feedforward network weight spaces. *Advances in Neural Computer*, *129–135*.

Hendrickx, J. M. H., & Flury, M. (2001). Uniform and preferential flow mechanisms in the vadose zone. In D. A. Feary (Ed.), *Conceptual models of flow and transport in the fractured vadose zone* (pp. 149–187). Washington, DC: National Research Council, National Academy Press.

Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B. M., Ribeiro, E., … Ruiperez Gonzalez, M. (2014). SoilGrids1km-global soil information based on automated mapping. *PLoS One*, *9*, e105992. https://doi.org/10.1371/journal.pone.0105992

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., … Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning (B Bond-Lamberty, ed.). *PLoS One*, *12*(2), e0169748. https://doi.org/10.1371/journal.pone.0169748

Heumann, S., Böttcher, J., & Springob, G. (2003). Pedotransfer functions for the pool size of slowly mineralizable organic N in sandy arable soils. *Journal of Plant Nutrition and Soil Science*, *166*(3), 308–318. https://doi.org/10.1002/jpln.200390048

Heumann, S., Ringe, H., & Böttcher, J. (2011). Field-specific simulations of net N mineralization based on digitally available soil and weather data: II. Pedotransfer functions for the pool sizes. *Nutrient Cycling in Agroecosystems*, *91*(3), 339–350. https://doi.org/10.1007/s10705-011-9465-x

Hinsinger, P., Brauman, A., Devau, N., Gérard, F., Jourdan, C., Laclau, J. P., … Plassard, C. (2011). Acquisition of phosphorus and other poorly mobile nutrients by roots. Where do plant nutrition models fail? *Plant and Soil*, *348*(1-2), 29–61. https://doi.org/10.1007/s11104-011-0903-y

Hoffmann, H., Zhao, G., Asseng, S., Bindi, M., Biernath, C., Constantin, J., … Ewert, F. (2016). Impact of spatial soil and climate input data aggregation on regional yield simulations. *PLoS One*, *11*(4), e0151782. https://doi.org/10.1371/journal.pone.0151782

Hollis, J. M., Hannam, J., & Bellamy, P. H. (2012). Empirically derived pedotransfer functions for predicting bulk density in European soils. *European Journal of Soil Science*, *63*(1), 96–109. https://doi.org/10.1111/j.1365-2389.2011.01412.x

Horn, A. L., Reiher, W., During, R. A., & Goth, S. (2006). Efficiency of pedotransfer functions describing cadmium sorption in soils. *Water, Air, and Soil Pollution*, *170*(1-4), 229–247. https://doi.org/10.1007/s11270-005-9003-z

Horton, R., Wierenga, P. J., & Nielsen, D. R. (1983). Evaluation of methods for determining the apparent thermal diffusivity of soil near the Surface1. *Soil Science Society of America Journal*, *47*(1), 25–32. https://doi.org/10.2136/sssaj1983.03615995004700010005x

Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics*, *15*(3), 651–674. https://doi.org/10.1198/106186006X133933

Huang, C.-Y., Hendrix, P. F., Fahey, T. J., Bohlen, P. J., & Groffman, P. M. (2010). A simulation model to evaluate the impacts of invasive earthworms on soil carbon dynamics. *Ecological Modelling*, *221*(20), 2447–2457. https://doi.org/10.1016/j.ecolmodel.2010.06.023

Huang, S., Bartlett, P., & Arain, M. A. (2016). An analysis of global terrestrial carbon, water and energy dynamics using the carbon-nitrogen coupled CLASS-CTEMN+ model. *Ecological Modelling*, *336*, 36–56. https://doi.org/10.1016/j.ecolmodel.2016.05.019

Hunt, E. R. Jr., Piper, S. C., Nemani, R., Keeling, C. D., Otto, R. D., & Running, S. W. (1996). Global net carbon exchange and intra-annual atmospheric CO_2 concentrations predicted by an ecosystem process model and three-dimensional atmospheric transport model. *Global Biogeochemical Cycles*, *10*(3), 431–456. https://doi.org/10.1029/96GB01691

Ippisch, O., Vogel, H. J., & Bastian, P. (2006). Validity limits for the van Genuchten–Mualem model and implications for parameter estimation and numerical simulation. *Advances in Water Resources*, *29*(12), 1780–1789. https://doi.org/10.1016/j.advwatres.2005.12.011

Jackson, T., Bindlish, R., & Zhao, T. (2011). *Vegetation index climatology*. Pasadena, CA: SMAP Science Document, Jet Propulsion Laboratory.

Jagtap, S. S., Lall, U., Jones, J. W., Gijsman, A. J., & Ritchie, J. T. (2004). Dynamic nearest-neighbor method for estimating soil water parameters. *Transactions of ASAE*, *47*(5), 1437–1444. https://doi.org/10.13031/2013.17623

Jalabert, S. S. M., Martin, M. P., Renaud, J. P., Boulonne, L., Jolivet, C., Montanarella, L., & Arrouays, D. (2010). Estimating forest soil bulk density using boosted regression modelling. *Soil Use and Management*, *26*(4), 516–528. https://doi.org/10.1111/j.1475-2743.2010.00305.x

Jame, Y. W., & Norum, D. I. (1980). Heat and mass transfer in a freezing unsaturated porous medium. *Water Resources Research*, *117*, 811–819.

Jana, R. B., & Mohanty, B. P. (2011). Enhancing PTFs with remotely sensed data for multi-scale soil water retention estimation. *Journal of Hydrology*, *399*(3-4), 201–211. https://doi.org/10.1016/j.jhydrol.2010.12.043

Jana, R. B., Mohanty, B. P., & Springer, E. P. (2007). Multiscale pedotransfer functions for soil water retention. *Vadose Zone Journal*, *6*(4), 868–878. https://doi.org/10.2136/vzj2007.0055

Jarvis, N. J. (2007). A review of non-equilibrium water flow and solute transport in soil macropores: Principles, controlling factors and consequences for water quality. *European Journal of Soil Science*, *58*(3), 523–546. https://doi.org/10.1111/j.1365-2389.2007.00915.x

Jarvis, N. (2008). Near-saturated hydraulic properties of macroporous soils. *Vadose Zone Journal*, *7*(4), 1302–1310. https://doi.org/10.2136/vzj2008.0065

Jarvis, N. J., Moeys, J., Hollis, J. M., Reichenberger, S., Lindahl, A. M. L., & Dubus, I. G. (2009). A conceptual model of soil susceptibility to macropore flow. *Vadose Zone Journal*, *8*(4), 902–910. https://doi.org/10.2136/vzj2008.0137

Jarvis, N., Koestel, J., Messing, I., Moeys, J., & Lindahl, A. (2013). Influence of soil, land use and climatic factors on the hydraulic conductivity of soil. *Hydrology and Earth System Sciences*, *17*(12), 5185–5195. https://doi.org/10.5194/hess-17-5185-2013

Javaux, M., Couvreur, V., Vanderborght, J., & Vereecken, H. (2013). Root water uptake: From three-dimensional biophysical processes to macroscopic modeling approaches. *Vadose Zone Journal*, *12*(4). https://doi.org/10.2136/vzj2013.02.0042

Jenny, H. (1941). Factors of soil formation: A system of quantitative pedology, pp. 1–261, McGraw Hill, data. *Journal of Petroleum Technology*, *5*(03), 71–78. https://doi.org/10.2118/225-G

Jing, X., Sanders, N. J., Shi, Y., Chu, H., Classen, A. T., Zhao, K., … He, J.-S. (2015). The links between ecosystem multifunctionality and above- and belowground biodiversity are mediated by climate. *Nature Communications*, *6*, 8159. https://doi.org/10.1038/ncomms9159

Johansen, O. (1975). Thermal conductivity of soils, PhD thesis, University of Trondheim, 236 pp., Universitetsbiblioteket i Trondheim, Høgskoleringen 1, 7034 Trondheim, Norway. Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/a044002.pdf

Jones, H. G. (2007). Monitoring plant and soil water status: Established and novel methods revisited and their relevance to studies of drought tolerance. *Journal of Experimental Botany*, *58*(2), 119–130. https://doi.org/10.1093/jxb/erl118

Jong, R. D., & Bootsma, A. (1996). Review of recent developments in soil water simulation models. *Canadian Journal of Soil Science*, *76*(3), 263–273. https://doi.org/10.4141/cjss96-033

Jorda, H., Bechtold, M., Jarvis, N., & Koestel, J. (2015). Using boosted regression trees to explore key factors controlling saturated and near-saturated hydraulic conductivity. *European Journal of Soil Science*, *66*(4), 744–756. https://doi.org/10.1111/ejss.12249

Kargas, G., Kerkides, P., Sotirakoglou, K., & Poulovassilis, A. (2016). Temporal variability of surface soil hydraulic properties under various tillage systems. *Soil and Tillage Research*, *158*, 22–31. https://doi.org/10.1016/j.still.2015.11.011

Karup, D., Moldrup, P., Paradelo, M., Katuwal, S., Norgaard, T., Greve, M. H., & de Jonge, L. W. (2016). Water and solute transport in agricultural soils predicted by volumetric clay and silt contents. *Journal of Contaminant Hydrology*, *192*, 194–202. https://doi.org/10.1016/j.jconhyd.2016.08.001

Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, *29*(2), 119–127. https://doi.org/10.2307/2986296

Kattge, J., Diaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Böniusch, G., … Wirth, C. (2011). TRY—A global database of plant traits. *Global Change Biology*, *17*(9), 2905–2935. https://doi.org/10.1111/j.1365-2486.2011.02451.x

Katuwal, S., Moldrup, P., Lamandé, M., Tuller, M., & de Jonge, L. W. (2015). Effects of CT number derived matrix density on preferential flow and transport in a macroporous agricultural soil. *Vadose Zone Journal*, *14*(7). https://doi.org/10.2136/vzj2015.01.0002

Kautz, T. (2014). Research on subsoil biopores and their functions in organically managed soils: A review. *Renewable Agriculture and Food Systems*, 1–10. https://doi.org/10.1017/S1742170513000549

Kerkhoff, A. J., Fagan, W. F., Elser, J. J., & Enquist, B. J. (2006). Phylogenetic and growth form variation in the scaling of nitrogen and phosphorus in the seed plants. *American Naturalist*, *168*, 103–122.

Khlosi, M., Alhamdoosh, M., Douaik, A., Gabriels, D., & Cornelis, W. M. (2016). Enhanced pedotransfer functions with support vector machines to predict water retention of calcareous soil. *European Journal of Soil Science*, *67*(3), 276–284. https://doi.org/10.1111/ejss.12345

Khodaverdiloo, H., Homaee, M., van Genuchten, M. T., & Dashtaki, S. G. (2011). Deriving and validating pedotransfer functions for some calcareous soils. *Journal of Hydrology*, *399*(1-2), 93–99. https://doi.org/10.1016/j.jhydrol.2010.12.040

Khosravi Moshizi, A., Heshmati, G. A., & Salman Mahiny, A. R. (2015). Identifying carbon sequestration hotspots in semiarid rangelands (Case study: Baghbazm region of Bardsir city, Kerman province). *Journal of Rangeland Science*, *5*(4), 325–335.

Kishné, A. S., Yimam, Y. T., Morgan, C. L., & Dornblaser, B. C. (2017). Evaluation and improvement of the default soil hydraulic parameters for the Noah land surface model. *Geoderma*, *285*, 247–259. https://doi.org/10.1016/j.geoderma.2016.09.022

Kitanidis, P. K. (1994). The concept of the dilution index. *Water Resources Research*, *30*(7), 2011–2026. https://doi.org/10.1029/94WR00762

Kleidon, A. (2004). Global datasets and rooting zone depth inferred from inverse methods. *J Climate*, *17*, 2714–2722.

Kleidon, A., & Heimann, M. (1998). A method of determining rooting depth from a terrestrial biosphere model and its impacts on the global water and carbon cycle. *Global Change Biology*, *4*(3), 275–286. https://doi.org/10.1046/j.1365-2486.1998.00152.x

Klier, C., Gayler, S., Haberbosch, C., Ruser, R., Stenger, R., Flessa, H., & Priesack, E. (2011). Modeling nitrous oxide emissions from potato-cropped soil. *Vadose Zone Journal*, *10*(1), 184–194. https://doi.org/10.2136/vzj2009.0194

Knudby, C., & Carrera, J. (2005). On the relationship between indicators of geostatistical, flow and transport connectivity. *Advances in Water Resources*, *28*(4), 405–421. https://doi.org/10.1016/j.advwatres.2004.09.001

Kobal, M., Urbancic, M., Potocic, N., de Vos, B., & Simoncic, P. (2011). Pedotransfer functions for bulk density estimation of forest soils. *Sumarski list*, *1-2*, 19–27.

Kodesova, R., Kocarek, M., Kodes, V., Drabek, O., Kozak, J., & Hejtmankova, K. (2011). Pesticide adsorption in relation to soil properties and soil type distribution in regional scale. *Journal of Hazardous Materials*, *186*(1), 540–550. https://doi.org/10.1016/j.jhazmat.2010.11.040

Koestel, J., & Jorda, H. (2014). What determines the strength of preferential transport in undisturbed soil under steady-state flow? *Geoderma*, *217-218*, 144–160. https://doi.org/10.1016/j.geoderma.2013.11.009

Koestel, J., & Larsbo, M. (2014). Imaging and quantification of preferential solute transport in soil macropores. *Water Resources Research*, *50*(5), 4357–4378. https://doi.org/10.1002/2014WR015351

Koestel, J. K., Moeys, J., & Jarvis, N. J. (2011). Evaluation of nonparametric shape measures for solute breakthrough curves. *Vadose Zone Journal*, *10*(4), 1261–1275. https://doi.org/10.2136/vzj2011.0010

Koestel, J. K., Moeys, J., & Jarvis, N. J. (2012). Meta-analysis of the effects of soil properties, site factors and experimental conditions on solute transport. *Hydrology and Earth System Sciences*, *16*(6), 1647–1665. https://doi.org/10.5194/hess-16-1647-2012

Koestel, J. K., Norgaard, T., Luong, N. M., Vendelboe, A. L., Moldrup, P., Jarvis, N. J., … Wollesen de Jonge, L. (2013). Links between soil properties and steady-state solute transport through cultivated topsoil at the field scale. *Water Resources Research*, *49*(2), 790–807. https://doi.org/10.1002/wrcr.20079

Köhne, J. M., Köhne, S., & Simunek, J. (2009a). A review of model applications for structured soils: Water flow and tracer transport. *Journal of Contaminant Hydrology*, *104*(1-4), 4–35. https://doi.org/10.1016/j.jconhyd.2008.10.002

Köhne, J. M., Köhne, S., & Simunek, J. (2009b). A review of model applications for structured soils: Pesticide transport. *Journal of Contaminant Hydrology*, *104*(1-4), 36–60. https://doi.org/10.1016/j.jconhyd.2008.10.003

Kolenbrander, G. J. (1970). Calculation of parameters for evaluation of leaching of salts under field conditions, illustrated by nitrate. *Plant and Soil*, *32*(1-3), 439–453. https://doi.org/10.1007/BF01372882

Koorevaar, P., Menelik, G., & Dirksen, C. (1983). Elements of soil physics. In *Developments in soils science* (Vol. 13, 230 pp.). New York: Elsevier.

Koster, R. D., Suarez, M. J., Ducharne, A., Stieglitz, M., & Kumar, P. (2000). A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. *Journal of Geophysical Research*, *105*(D20), 24,809–24,822. https://doi.org/10.1029/2000JD900327

Kosugi, K. (1996). Lognormal distribution model for unsaturated soil hydraulic properties. *Water Resources Research*, *32*(9), 2697–2703. https://doi.org/10.1029/96WR01776

Kosugi, K., Hopmans, J. W., & Dane, J. H. (2002). Water retention and storage—Parametric models. In J. H. Dane & G. C. Topp (Eds.), *Methods of soil analysis. Part 4. Physical methods*, Soil Science Society of America Book Series (Vol. 5, pp. 739–758). Madison, WI: Oil Science Society of America.

Kowalczyk, E. A., Wang, Y. P., Law, R. M., Davies, H. L., McGregor, J. L., and Abramowitz, G. (2006). The CSIRO Atmosphere Biosphere Land Exchange (CABLE) model for use in climate models and as an offline model, CSIRO, ISBN 1921232390.

Kuhnert, M., Yeluripati, J., Smith, P., Hoffmann, H., van Oijen, M., Constantin, J., … Ewert, F. (2016). Impact analysis of climate data aggregation at different spatial scales on simulated net primary productivity for croplands. *European Journal of Agronomy*, *88*, 41–52. https://doi.org/10.1016/j.eja.2016.06.005

Kutílek, M. and Nielsen, D. R. (1994). Soil hydrology. Catena-Verlag.

Lal, R. (2005). Forest soils and carbon sequestration. *Forest Ecology and Management*, *220*(1-3), 242–258. https://doi.org/10.1016/j.foreco.2005.08.015

Lal, R., & Shukla, M. K. (2004). *Principles of soil physics*. New York: Marcel Dekker, Inc.

Lall, U., & Sharma, A. (1996). A nearest-neighbor bootstrap for resampling hydrologic time series. *Water Resources Research*, *32*(3), 679–693. https://doi.org/10.1029/95WR02966

Lamorski, K., Pachepsky, Y., Sławiński, C., & Walczak, R. T. (2008). Using support vector machines to develop pedotransfer functions for water retention of soils in Poland. *Soil Science Society of America Journal*, *72*(5), 1243–1247. https://doi.org/10.2136/sssaj2007.0280N

Larsbo, M., Roulier, S., Stenemo, F., Kasteel, R., & Jarvis, N. (2005). An improved dual-permeability model of water flow and solute transport in the vadose zone. *Vadose Zone Journal*, *4*(2), 398–406. https://doi.org/10.2136/vzj2004.0137

Larsbo, M., Koestel, J., & Jarvis, N. (2014). Relations between macropore network characteristics and the degree of preferential solute transport. *Hydrology and Earth System Sciences*, *18*(12), 5255–5269. https://doi.org/10.5194/hess-18-5255-2014

Larsbo, M., Koestel, J., Kätterer, T., & Jarvis, N. (2016). Preferential transport in macropores is reduced by soil organic carbon. *Vadose Zone Journal*, *15*(9). https://doi.org/10.2136/vzj2016.03.0021

Laudone, G. M., Matthews, G. P., Bird, N. R. A., Whalley, W. R., Cardenas, L. M., & Gregory, A. S. (2011). A model to predict the effects of soil structure on denitrification and N_2O emission. *Journal of Hydrology*, *409*(1-2), 283–290. https://doi.org/10.1016/j.jhydrol.2011.08.026

Lavorel, S., Diaz, S., Prentice, I. C., & Leadley, P. (2008). Refining plant functional classifications for Earth system modeling. *Global Land Project (GLP) Newsletter*, *3*, 38–40.

Lawrence, C. R., Neff, J. C., & Schimel, J. P. (2009). Does adding microbial mechanisms of decomposition improve soil organic matter models? A comparison of four models using data from a pulsed rewetting experiment. *Soil Biology and Biochemistry*, *41*(9), 1923–1934. https://doi.org/10.1016/j.soilbio.2009.06.016

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., … Slater, A. G. (2011). Parameterization improvements and functional and structural advances in version 4 of the community land model. *Journal of Advances in Modeling Earth Systems*, *3*(3), M03001. https://doi.org/10.1029/2011MS000045

Lee, T. J. (1992). The impact of vegetation on the atmospheric boundary layer and convective storms. Atmospheric science paper no. 509. Fort Collins, CO: Department of Atmospheric Science, Colorado State University.

Lee, T. J., & Pielke, R. A. (1992). Estimating the soil surface specific humidity. *Journal of Applied Meteorology*, *31*(5), 480–484. https://doi.org/10.1175/1520-0450(1992)031%3C0480:ETSSSH%3E2.0.CO;2

Lefcheck, J. S., Byrnes, J. E. K., Isbell, F., Gamfeldt, L., Griffin, J. N., Eisenhauer, N., … Duffy, J. E. (2015). Biodiversity enhances ecosystem multifunctionality across trophic levels and habitats. *Nature Communications*, *6*, 6936. https://doi.org/10.1038/ncomms7936

Leij, F. J., Alves, W. J., van Genuchten, M. T., and Williams, J. R. (1996). Unsaturated soil hydraulic database, UNSODA 1.0 User's manual. Report EPA/600/R-96/095 (103 pp.). Ada, OK: U.S. Environmental Protection Agency.

Leij, F. J., Ghezzehei, T. A., & Or, D. (2002). Modeling the dynamics of the soil pore-size distribution. *Soil and Tillage Research*, *64*(1-2), 61–78. https://doi.org/10.1016/S0167-1987(01)00257-4

Leitner, D., Klepsch, S., Bodner, G., & Schnepf, A. (2010). A dynamic root system growth model based on L-systems. *Plant and Soil*, *332*(1-2), 177–192. https://doi.org/10.1007/s11104-010-0284-7

Letts, M. G., Roulet, N. T., Corner, N. T., Skapura, M. R., & Verseghy, D. L. (2000). Parametrization of peatland hydraulic properties for the Canadian land surface scheme. *Atmosphere-Ocean*, *38*(1), 141–160.

Levi, M. R., Schaap, M. G., & Rasmussen, C. (2015). Application of spatial pedotransfer functions to understand soil modulation of vegetation response to climate. *Vadose Zone Journal*, *14*(9). https://doi.org/10.2136/vzj2014.09.0126

Li, X., Li, J. H., & Zhang, L. M. (2014). Predicting bimodal soil—Water characteristic curves and permeability functions using physically based parameters. *Computers and Geotechnics*, *57*, 85–96. https://doi.org/10.1016/j.compgeo.2014.01.004

Li, L., Maher, K., Navarre-Sitchler, A., Druhan, J., Meile, C., Lawrence, C., … Beisman, J. (2017). Expanding the role of reactive transport models in critical zone processes. *Earth-Science Reviews*, *165*, 280–301. https://doi.org/10.1016/j.earscirev.2016.09.001

Liao, K., Xu, S., & Zhu, Q. (2015). Development of ensemble pedotransfer functions for cation exchange capacity of soils of Qingdao in China. *Soil Use and Management*, *31*(4), 483–490. https://doi.org/10.1111/sum.12207

Lichner, L., Hallett, P. D., Drongova, Z., Czachor, H., Kovacik, L., Mataix-Solera, J., & Homolák, M. (2013). Algae influence the hydrophysical parameters of a sandy soil. *Catena*, *108*, 58–68. https://doi.org/10.1016/j.catena.2012.02.016

Lilly, A., Nemes, A., Rawls, W. J., & Pachepsky, Y. A. (2008). Probabilistic approach to the identification of input variables to estimate hydraulic conductivity. *Soil Science Society of America Journal*, *72*(1), 16–24. https://doi.org/10.2136/sssaj2006.0391

Lin, H. S., McInnes, K. J., Wilding, L. P., & Hallmark, C. T. (1999). Effects of soil morphology on hydraulic properties: II. Hydraulic pedotransfer functions. *Soil Science Society of America Journal*, *63*(4), 955–961. https://doi.org/10.2136/sssaj1999.634955x

Lloyd, J., & Taylor, J. A. (1994). On the temperature dependence of soil respiration. *Functional Ecology*, *8*(3), 315–323. https://doi.org/10.2307/2389824

Long, S. P. (1991). Modification of the response of photosynthetic productivity to rising temperature by atmospheric CO_2 concentrations: Has its importance been underestimated? *Plant, Cell & Environment*, *14*(8), 729–739. https://doi.org/10.1111/j.1365-3040.1991.tb01439.x

Lu, S., Ren, T., Gong, Y., & Horton, R. (2007). An improved model for predicting soil thermal conductivity from water content at room temperature. *Soil Science Society of America Journal*, *71*(1), 8–14. https://doi.org/10.2136/sssaj2006.0041

Ludwig, B., Schulz, E., Rethemeyer, J., Merbach, I., & Flessa, H. (2007). Predictive modelling of C dynamics in the long-term fertilization experiment at Bad Lauchstädt with the Rothamsted carbon model. *European Journal of Soil Science*, *58*(5), 1155–1163. https://doi.org/10.1111/j.1365-2389.2007.00907.x

Luo, Y., Ahlström, A., Allison, S. D., Batjes, N. H., Brovkin, V., Carvalhais, N., … Zhou, T. (2016). Toward more realistic projections of soil carbon dynamics by Earth system models. *Global Biogeochemical Cycles*, *30*, 40–56. https://doi.org/10.1002/2015GB005239

Maestre, F. T., Delgado-Baquerizo, M., Jeffries, T. C., Eldridge, D. J., Ochoa, V., Gozalo, B., … Singh, B. K. (2015). Increasing aridity reduces soil microbial diversity and abundance in global drylands. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(51), 15,684–15,689. https://doi.org/10.1073/pnas.1516684112

Maire, V., Wright, I. J., Prentice, I. C., Batjes, N. H., Bhaskar, R., van Bodegom, P. M., … Santiago, L. S. (2015). Global effects of soil and climate on leaf photosynthetic traits and rates. *Global Ecology and Biogeography*, *24*(6), 706–717. https://doi.org/10.1111/geb.12296

Malavasi, U. C., Davis, A. S., & Malavasi, M. M. (2016). Estimating water in living woody stems - a review. *Cerne*, *22*(4), 415–422. https://doi.org/10.1590/01047760201622032169

Manrique, L. A., & Jones, C. A. (1991). Bulk density of soils in relation to soil physical and chemical properties. *Soil Science Society of America Journal*, *55*(2), 476–481. https://doi.org/10.2136/sssaj1991.03615995005500020030x

Manzoni, S., & Porporato, A. (2009). Soil carbon and nitrogen mineralization: Theory and models across scales. *Soil Biology and Biochemistry*, *41*(7), 1355–1379. https://doi.org/10.1016/j.soilbio.2009.02.031

Manzoni, S., Schimel, J. P., & Porporato, A. (2012). Responses of soil microbial communities to water stress: Results from a meta-analysis. *Ecology*, *93*(4), 930–938. https://doi.org/10.1890/11-0026.1

Maren, A. J., Harston, C. T., & Pap, R. M. (2014). *Handbook of neural computing applications*. San Diego, CA: Academic Press.

Marthews, T. R., Quesada, C. A., Galbraith, D. R., Malhi, Y., Mullins, C. E., Hodnett, M. G., & Dharssi, I. (2014). High-resolution hydraulic parameter maps for surface soils in tropical South America. *Geoscientific Model Development*, *7*(3), 711–723. https://doi.org/10.5194/gmd-7-711-2014

Martín, M. Á., Reyes, M., & Taguas, F. J. (2017). Estimating soil bulk density with information metrics of soil texture. *Geoderma*, *287*, 66–70. https://doi.org/10.1016/j.geoderma.2016.09.008

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., … Voldoire, A. (2013). The SURFEXv7.2 externalized platform for the simulation of Earth surface variables and fluxes. *Geoscientific Model Development*, *6*(4), 929–960. https://doi.org/10.5194/gmd-6-929-2013

Mayr, T., & Jarvis, N. J. (1999). Pedotransfer functions to estimate soil water retention parameters for a modified Brooks–Corey type model. *Geoderma*, *91*(1-2), 1–9. https://doi.org/10.1016/S0016-7061(98)00129-3

McBratney, A. B., Minasny, B., Cattle, S. R., & Vervoort, R. W. (2001). From pedotransfer functions to soil inference systems. *Geoderma*, *109*, 41–73.

McBratney, A. B., Minasny, B., Cattle, S. R., & Vervoort, R. (2002). From pedotransfer functions to soil inference systems. *Geoderma*, *109*(1-2), 41–73. https://doi.org/10.1016/S0016-7061(02)00139-8

McBratney, A. B., Minasny, B., & Rossel, R. V. (2006). Spectral soil analysis and inference systems: A powerful combination for solving the soil data crisis. *Geoderma*, *136*(1-2), 272–278. https://doi.org/10.1016/j.geoderma.2006.03.051

McBratney, A. B., Minasny, B., & Tranter, G. (2011). Necessary meta-data for pedotransfer functions. *Geoderma*, *160*(3-4), 627–629. https://doi.org/10.1016/j.geoderma.2010.09.023

McClain, M. E., Boyer, E. W., Dent, C. L., Gergel, S. E., Grimm, N. B., Groffman, P. M., … Pinay, G. (2003). Biogeochemical hot spots and hot moments at the Interface of terrestrial and aquatic ecosystems. *Ecosystems*, *6*(4), 301–312. https://doi.org/10.1007/s10021-003-0161-9

McCumber, M. C., & Pielke, R. A. (1981). Simulation of the effects of surface fluxes of heat and moisture in a mesoscale numerical model: 1. Soil layer. *Journal of Geophysical Research*, *86*(C10), 9929–9938. https://doi.org/10.1029/JC086iC10p09929

McKenzie, N., & Jacquier, D. (1997). Improving the field estimation of saturated hydraulic conductivity in soil survey. *Soil Research*, *35*(4), 803–827. https://doi.org/10.1071/S96093

McKenzie, N. J., & Ryan, P. J. (1999). Spatial prediction of soil properties using environmental correlation. *Geoderma*, *89*(1-2), 67–94. https://doi.org/10.1016/S0016-7061(98)00137-2

McMahon, S. M., Harrison, S. P., Armbruster, W. S., Bartlein, P. J., Beale, C. M., Edwards, M. E., … Prentice, I. C. (2011). Improving assessment and modelling of climate change impacts on global terrestrial biodiversity. *Trends in Ecology & Evolution*, *26*(5), 249–259. https://doi.org/10.1016/j.tree.2011.02.012

Meyer, P. D., & Gee, G. W. (1999). Flux-based estimation of field capacity. *Journal of Geotechnical and Geoenviromental Engineering*, *125*(7), 595–599. https://doi.org/10.1061/(ASCE)1090-0241(1999)125:7(595)

Millington, R. J., & Quirk, J. P. (1959). Permeability of porous media. *Nature*, *183*(4658), 387–388. https://doi.org/10.1038/183387a0

Minasny, B., & McBratney, A. (2002). The neuro-m method for fitting neural network parametric pedotransfer functions. *Soil Science Society of America Journal*, *66*(2), 352–361.

Minasny, B., & Perfect, E. (2004). Solute adsorption and transport parameters. *Development of Pedotransfer Functions in Soil Hydrology*, *30*, 195–224. https://doi.org/10.1016/s0166-2481(04)30012-7

Minasny, B., McBratney, A. B., & Bristow, K. L. (1999). Comparison of different approaches to the development of pedotransfer functions for water-retention curves. *Geoderma*, *93*(3-4), 225–253. https://doi.org/10.1016/S0016-7061(99)00061-0

Minasny, B., Hopmans, J. W., Harter, T., Eching, S. O., Tuli, A., & Denton, M. A. (2004). Neural networks prediction of soil hydraulic functions for alluvial soils using multistep outflow data. *Soil Science Society of America Journal*, *68*(2), 417–429. https://doi.org/10.2136/sssaj2004.4170

Minasny, B., McBratney, A. B., Lendoca-Santos, M. L., Odeh, I. O. A., & Guyon, B. (2006). Prediction and digital mapping of soil carbon storage in the lower Namoi Valley. *Australian Journal of Soil Research*, *44*(3), 233–244. https://doi.org/10.1071/SR05136

Minasny, B., Finke, P., Stockmann, U., Vanwalleghem, T., & McBratney, A. B. (2015). Resolving the integral connection between pedogenesis and landscape evolution. *Earth-Science Reviews*, *150*, 102–120.

Mishra, U., & Riley, W. J. (2012). Alaskan soil carbon stocks: Spatial variability and dependence on environmental factors. *Biogeosciences*, *9*(9), 3637–3645. https://doi.org/10.5194/bg-9-3637-2012

Mishra, U., & Riley, W. J. (2014). Active-layer thickness across Alaska: Comparing observation-based estimates with CMIP5 Earth system model predictions. *Soil Science Society of America Journal*, *78*(3), 894–902. https://doi.org/10.2136/sssaj2013.11.0484

Mishra, U., & Riley, W. J. (2015). Scaling impacts on environmental controls and spatial heterogeneity of soil organic carbon stocks. *Biogeosciences*, *12*(13), 3993–4004. https://doi.org/10.5194/bg-12-3993-2015

Mishra, U., Drewniak, B., Jastrow, J. D., Matamala, R. M., & Vitharana, U. W. A. (2017). Spatial representation of high latitude organic carbon and active-layer thickness in CMIP5 Earth system models. *Geoderma*, *300*, 55–63. https://doi.org/10.1016/j.geoderma.2016.04.017

Mittermeier, R. A., Turner, W. R., Larsen, F. W., Brooks, T. M., & Gascon, C. (2011). Global biodiversity conservation: The critical role of hotspots. In F. E. Zachos & J. C. Habel (Eds.), *Biodiversity hotspots* (pp. 3–22). London: Springer Publishers. https://doi.org/10.1007/978-3-642-20992-5_1

Moeys, J., Bergheaud, V., & Coquet, Y. (2011). Pedotransfer functions for isoproturon sorption on soils and vadose zone materials. *Pest Management Science*, *67*(10), 1309–1319. https://doi.org/10.1002/ps.2187

Moeys, J., Larsbo, M., Bergström, L., Brown, C. D., Coquet, Y., & Jarvis, N. J. (2012). Functional test of pedotransfer functions to predict water flow and solute transport with the dual-permeability model MACRO. *Hydrology and Earth System Sciences*, *16*, 2069–2083. https://doi.org/10.5194/hess-16-2069-2012

Mohanty, B. P. (2013). Soil hydraulic property estimation using remote sensing: A review. *Vadose Zone Journal*, *12*(4). https://doi.org/10.2136/vzj2013.06.0100

Moldrup, P., Olesen, T., Komatsu, T., Yoshikawa, S., Schjonning, P., & Rolston, D. E. (2003). Modeling diffusion and reaction in soils: X. A unifying model for solute and gas diffusivity in unsaturated soil. *Soil Science*, *168*(5), 321–337. https://doi.org/10.1097/00010694-200305000-00002

Montzka, C., Moradkhani, H., Weihermueller, L., Franssen, H.-J. H., Canty, M., & Vereecken, H. (2011). Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter. *Journal of Hydrology*, *399*(3-4), 410–421. https://doi.org/10.1016/j.jhydrol.2011.01.020

Montzka, C., Herbst, M., Weihermüller, L., Verhoef, A., & Vereecken, H. (2017a). A global data set of soil hydraulic properties and sub-grid variability of soil water retention and hydraulic conductivity curves. *Earth System Science Data*, *9*, 529–543. https://doi.org/10.5194/essd-2017-13

Montzka, C., Herbst, M., Weihermüller, L., Verhoef, A., & Vereecken, H. (2017b). A global data set of soil hydraulic properties and sub-grid variability of soil 2012. A pseudo-continuous neural network approach for developing water retention pedotransfer functions with limited data. *Journal of Hydrology*, *442-443*, 46–54. https://doi.org/10.1016/j.jhydrol.2012.03.036

Moore, T. R., De Souza, W., & Koprivnjak, J.-F. (1992). Controls on the sorption of dissolved organic carbon by soils. *Soil Science*, *154*(2), 120–129. https://doi.org/10.1097/00010694-199208000-00005

Morris, J. C. (2015). General method for predicting soil data via pattern matching on pedotransfer functions.

Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(3), 247–259. https://doi.org/10.1016/j.isprsjprs.2010.11.001

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

Mualem, Y., & Dagan, G. (1978). Hydraulic conductivity of soils: Unified approach to the statistical models. *Soil Science Society of America Journal*, *42*(3), 392–395. https://doi.org/10.2136/sssaj1978.03615995004200030003x

Mulla, D. J. (2012). Chapter 20—Modeling and mapping soil spatial and temporal variability A2 - Lin, Henry. In *Hydropedology* (pp. 637–664). Boston: Academic Press. https://doi.org/10.1016/B978-0-12-386941-8.00020-4

Nachabe, M. H. (1998). Refining the definition of field capacity in the literature. *Journal of Irrigation and Drainage Engineering*, *124*(4), 230–232. https://doi.org/10.1061/(ASCE)0733-9437(1998)124:4(230)

Nachtergaele, F., van Velthuize, H., Verelst, L., Wiberg, D., Batjes, N., Dijkshoorn, K., & Shi, X. (2012). Harmonized world soil database, version 1.2, FAO/IIASA/ISRIC/ISS-CAS/JRC, FAO, Rome, Italy and IIASA, Laxenburg, Austria. Retrieved from http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HWSD_Documentation.pdf

Naipal, V., Reick, C., Pongratz, J., & Van Oost, K. (2015). Improving the global applicability of the RUSLE model—Adjustment of the topographical and rainfall erosivity factors. *Geoscientific Model Development*, *8*(9), 2893–2913. https://doi.org/10.5194/gmd-8-2893-2015

Nanko, K., Ugawa, S., Hashimoto, S., Imaya, A., Kobayashi, M., Sakai, H., … Kaneko, S. (2014). A pedotransfer function for estimating bulk density of forest soil in Japan affected by volcanic ash. *Geoderma*, *213*, 36–45. https://doi.org/10.1016/j.geoderma.2013.07.025

National Cooperative Soil Survey (2017). National Cooperative Soil Survey Characterization Database. Retrieved from http://ncsslabdatamart.sc.egov.usda.gov/, Accessed 2017–01-01.

Nemes, A. (2011). Databases of soil physical and hydraulic properties. Encyclopedia of agrophysics. *Encyclopedia of Earth Sciences Series*, *2011*(Part 4), 194–199. https://doi.org/10.1007/978-90-481-3585-1_39

Nemes, A. (2015). Why do they keep rejecting my manuscript—Do's and don'ts and new horizons in pedotransfer studies.Agrokémia és. *Talajtan*, *64*(2), 361–371. https://doi.org/10.1556/0088.2015.64.2.4

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*(3-4), 151–162. https://doi.org/10.1016/S0022-1694(01)00465-6

Nemes, A., Schaap, M. G., & Wösten, J. H. M. (2003). Functional evaluation of pedotransfer functions derived from different scales of data collection. *Soil Science Society of America Journal*, *67*(4), 1093–1102. https://doi.org/10.2136/sssaj2003.1093

Nemes, A., Rawls, W. J., & Pachepsky, Y. A. (2005). Influence of organic matter on the estimation of saturated hydraulic conductivity. *Soil Science Society of America Journal*, *69*(4), 1330–1337. https://doi.org/10.2136/sssaj2004.0055

Nemes, A., Rawls, W. J., & Pachepsky, Y. A. (2006). Use of the nonparametric nearest neighbor approach to estimate soil hydraulic properties. *Soil Science Society of America Journal*, *70*(2), 327–336. https://doi.org/10.2136/sssaj2005.0128

Nemes, A., Rawls, W. J., Pachepsky, Y. A., & Van Genuchten, M. T. (2006). Sensitivity analysis of the nonparametric nearest neighbor technique to estimate soil water retention. *Vadose Zone Journal*, *5*(4), 1222–1235. https://doi.org/10.2136/vzj2006.0017

Nemes, A., Timlin, D. J., Pachepsky, Y. A., & Rawls, W. J. (2009). Evaluation of the Rawls et al. (1982) pedotransfer functions for their applicability at the U.S. national scale. *Soil Science Society of America Journal*, *73*(5), 1638–1645. https://doi.org/10.2136/sssaj2008.0298

Nemes, A., Timlin, D. J., & Quebedeaux, B. (2010). Ensemble approach to provide uncertainty estimates of soil bulk density in support of simulation-based environmental risk assessment studies. *Soil Science Society of America Journal*, *74*(6), 1938–1945. https://doi.org/10.2136/sssaj2009.0370

Nemes, A., Pachepsky, Y. A., & Timlin, D. J. (2011). Toward improving global estimates of field soil water capacity. *Soil Science Society of America Journal*, *75*(3), 807–812. https://doi.org/10.2136/sssaj2010.0251

Nguyen, P. M., De Pue, J., Van Le, K., & Cornelis, W. (2015). Impact of regression methods on improved effects of soil structure on soil water retention estimates. *Journal of Hydrology*, *525*, 598–606. https://doi.org/10.1016/j.jhydrol.2015.04.014

Niemann, W. L., & Rovey, C. W. II (2009). A systematic field-based testing program of hydraulic conductivity and dispersivity over a range in scale. *Hydrogeology Journal*, *17*(2), 307–320. https://doi.org/10.1007/s10040-008-0365-3

Niu, G.-Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., … Xia, Y. (2011). The community Noah land surface model with multi-parameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research*, *116*, D12109. https://doi.org/10.1029/2010JD015139

Niu, S., Classen, A. T., Dukes, J. S., Kardol, P., Liu, L., Luo, Y., … Zaehle, S. (2016). Global patterns and substrate-based mechanisms of the terrestrial nitrogen cycle. *Ecology Letters*, *19*(6), 697–709. https://doi.org/10.1111/ele.12591

Nkedi-Kizza, P., Biggar, J. W., Selim, H. M., Vangenuchten, M. T., Wierenga, P. J., Davidson, J. M., & Nielsen, D. R. (1984). On the equivalence of two conceptual models for describing ion-exchange during transport through an aggregated oxisoil. *Water Resources Research*, *20*(8), 1123–1130. https://doi.org/10.1029/WR020i008p01123

Noilhan, J., & Lacarrère, P. (1991). GCM gridscale evaporation from mesoscale modelling. In *ECMWF Proceedings* (pp. 245–274). Reading, UK.

Nouri, H., Beecham, S., Anderson, S., & Nagler, P. (2014). High spatial resolution WorldView-2 imagery for mapping NDVI and its relationship to temporal urban landscape evapotranspiration factors. *Remote Sensing*, 6(1), 580–602. https://doi.org/10.3390/rs6010580

Nussbaum, M., Papritz, A., Zimmermann, S., & Walthert, L. (2016). Pedotransfer function to predict density of forest soils in Switzerland. *Journal of Plant Nutrition and Soil Science*, 179(3), 321–326. https://doi.org/10.1002/jpln.201500546

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., … Yang, Z.-L. (2013). Technical description of version 4.5 of the community land model (CLM). NCAR Technical Note NCAR/TN-503+STR (422 pp.). Boulder, CO: National Center for Atmospheric Research. https://doi.org/10.5065/D6RR1W7M

Oosterveld, M., & Chang, C. (1980). Empirical relationship between laboratory determinations of soil texture and moisture retention. *Canadian Agricultural Engineering*, 22, 149–151.

Oosterwoud, M. R., van der Ploeg, M. J., van der Schaaf, S., & van der Zee, S. E. A. T. M. (2017). Variation in hydrologic connectivity as a result of microtopography explained by discharge to catchment size relationship. *Hydrological Processes*, 31(15), 2683–2699. https://doi.org/10.1002/hyp.11164

Over, M. W., Wollschläger, U., Osorio-Murillo, C. A., & Rubin, Y. (2015). Bayesian inversion of Mualem-van Genuchten parameters in a multi-layer soil profile: A data-driven, assumption-free likelihood function. *Water Resources Research*, 51(2), 861–884. https://doi.org/10.1002/2014WR015252

Pachepsky, Y., & Hill, R. L. (2017). Scale and scaling in soils. *Geoderma*, 287, 4–30. https://doi.org/10.1016/j.geoderma.2016.08.017

Pachepsky, Y. A., & Rawls, W. J. (1999). Accuracy and reliability of pedotransfer functions as affected by grouping soils. *Soil Science Society of America Journal*, 63(6), 1748. https://doi.org/10.2136/sssaj1999.6361748x

Pachepsky, Y. A., & Rawls, W. J. (2003). Soil structure and pedotransfer functions. *European Journal of Soil Science*, 54(3), 443–452. https://doi.org/10.1046/j.1365-2389.2003.00485.x

Pachepsky, Y. A., & Rawls, W. J. (2004). *Development of pedotransfer functions in soil hydrology*. Elsevier, Amsterdam: Developments in Soil Science.

Pachepsky, Y. A., Shcherbakov, R. A., Várallyay, G., & Rajkai, K. (1982). Soil water retention as related to other soil physical properties. *Pochvovedenie*, 2, 42–52.

Pachepsky, Y. A., Timlin, D., & Varallyay, G. (1996). Artificial neural networks to estimate soil water retention from easily measurable data. *Soil Science Society of America Journal*, 60(3), 727. https://doi.org/10.2136/sssaj1996.03615995006000030007x

Pachepsky, Y. A., Rawls, W. J., & Lin, H. S. (2006). Hydropedology and pedotransfer functions. *Geoderma*, 131(3-4), 308–316. https://doi.org/10.1016/j.geoderma.2005.03.012

Pages, L., Vercambre, G., Drouet, J. L., Lecompte, F., Collet, C., & Le Bot, J. (2004). Root Typ: A generic model to depict and analyse the root system architecture. *Plant and Soil*, 258(1), 103–119. https://doi.org/10.1023/B:PLSO.0000016540.47134.03

Paloscia, S., Pettinato, S., Santi, E., Notarnicola, C., Pasolli, L., & Reppucci, A. (2013). Soil moisture mapping using Sentinel-1 images: Algorithm lk preliminary validation. *Remote Sensing of Environment*, 134, 234–248. https://doi.org/10.1016/j.rse.2013.02.027

Palta, M. M., Ehrenfeld, J. G., & Groffman, P. M. (2014). "Hotspots" and "hot moments" of denitrification in urban brownfield wetlands. *Ecosystems*, 17(7), 1121–1137. https://doi.org/10.1007/s10021-014-9778-0

Paradelo, M., Katuwal, S., Moldrup, P., Norgaard, T., Herath, L., & de Jonge, L. W. (2016). X-ray CT-derived soil characteristics explain varying air, water, and solute transport properties across a loamy field. *Vadose Zone Journal*, 15(4). https://doi.org/10.2136/vzj2015.07.0104

Parasuraman, K., Elshorbagy, A., & Si, B. C. (2006). Estimating saturated hydraulic conductivity in spatially variable fields using neural network ensembles. *Soil Science Society of America Journal*, 70(6), 1851–1859. https://doi.org/10.2136/sssaj2006.0045

Patil, N. G., & Chaturvedi, A. (2012). Estimation of bulk density of waterlogged soils from basic properties. *Archives of Agronomy and Soil Science*, 58(5), 499–509. https://doi.org/10.1080/03650340.2010.530254

Patil, N. G., Pal, D. K., Mandal, C., & Mandal, D. K. (2012). Soil water retention characteristics of vertisols and pedotransfer functions based on nearest neighbor and neural networks approaches to estimate AWC. *Journal of Irrigation and Drainage Engineering*, 138(2), 177–184. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000375

Patterson, K. A. (1990). Global Distributions of Total and Total-Avaiable Soil Water-Holding Capacities. University of Delaware.

Paydar, Z., & Cresswell, H. P. (1996). Water retention in Australian soils. Prediction using particle size, bulk density, and other properties. *Soil Research*, 34(5), 679–693. https://doi.org/10.1071/SR9960679

Pelletier, J. D., Barron-Gafford, G. A., Breshears, D. D., Brooks, P. D., Chorover, J., Durcik, M., & Troch, P. A. (2013). Coevolution of nonlinear trends in vegetation, soils, and topography with elevation and slope aspect: A case study in the sky islands of southern Arizona. *Journal of Geophysical Research: Earth Surface*, 118, 741–758. https://doi.org/10.1002/jgrf.20046

Perfect, E. (2003). A pedotransfer function for predicting solute dispersivity: Model testing and upscaling. In D. E. Radcliffe, H. M. Selim, & Y. Pachepsky (Eds.), *Scaling Methods in Soil Physics* (1st ed., 456 pp.). Boca Raton, FL: CRC Press. https://doi.org/10.1201/9780203011065.ch6

Perfect, E., Sukop, M. C., & Haszler, G. R. (2002). Prediction of dispersivity for undisturbed soil columns from water retention parameters. *Soil Science Society of America Journal*, 66(3), 696–701. https://doi.org/10.2136/sssaj2002.6960

Peters, A., & Durner, W. (2008). A simple model for describing hydraulic conductivity in unsaturated porous media accounting for film and capillary flow. *Water Resources Research*, 44(11), W11417. https://doi.org/10.1029/2008WR007136

Peters-Lidard, C. D., Blackburn, E., Liang, X., & Wood, E. F. (1998). The effect of soil thermal conductivity parameterization on surface energy fluxes and temperatures. *Journal of the Atmospheric Sciences*, 55(7), 1209–1224. https://doi.org/10.1175/1520-0469(1998)055%3C1209:TEOSTC%3E2.0.CO;2

Pitman, A. J. (2003). The evolution of, and revolution in, land surface schemes designed for climate models. *International Journal of Climatology*, 23(5), 479–510. https://doi.org/10.1002/joc.893

Press, W., Teukolsky, S., Vetterling, W., & Flannery, B. (1992). *Numerical recipes in C: The art of scientific computing*. Cambridge: Cambridge University Press.

Priesack, E., Gayler, S., & Hartmann, H. P. (2006). The impact of crop growth sub-model choice on simulated water and nitrogen balances. *Nutrient Cycling in Agroecosystems*, 75(1-3), 1–13. https://doi.org/10.1007/s10705-006-9006-1

Priesack, E., Berkenkamp, A., Gayler, S., Hartmann, H. P., & Loos, C. (2008). Development and application of agro-ecosystem models, 329-349.

Pringle, M. J., Romano, N., Minasny, B., Chirico, G. B., & Lark, R. M. (2007). Spatial evaluation of pedotransfer functions using wavelet analysis. *Journal of Hydrology*, 333(2-4), 182–198. https://doi.org/10.1016/j.jhydrol.2006.08.007

Purves, D., & Pacala, S. (2008). Predictive models of forest dynamics. *Science*, 320(5882), 1452–1453. https://doi.org/10.1126/science.1155359

Qiu, J. (2014). Land models put to climate test. *Nature*, 510(7503), 16–17. https://doi.org/10.1038/510016a

Ramirez, K., Döring, M., Eisenhauer, N., Gardi, C., Ladau, J., Leff, J. W., … Wall, D. H. (2015). Toward a global platform for linking soil biodiversity data. *Frontiers in Ecology and Evolution*, 3(91). https://doi.org/10.3389/fevo.2015.00091

Ramírez, B. H., van der Ploeg, M., Teuling, A. J., Ganzeveld, L., & Leemans, R. (2017). Tropical montane cloud forests in the Orinoco river basin: The role of soil organic layers in water storage and release. *Geoderma*, *298*, 14–26. https://doi.org/10.1016/j.geoderma.2017.03.007

Rasiah, V. (1995). Comparison of pedotransfer functions to predict nitrogen-mineralization parameters of one- and two-pool models. *Communications in Soil Science and Plant Analysis*, *26*(11-12), 1873–1884. https://doi.org/10.1080/00103629509369414

Rasiah, V., & Kay, B. D. (1998). Legume N mineralization: Effect of aeration and size distribution of water-filled pores. *Soil Biology and Biochemistry*, *30*(1), 89–96. https://doi.org/10.1016/S0038-0717(97)00083-7

Rasmussen, C. E. (2004). Gaussian processes in machine learning. Advanced lectures on machine learning. *Lecture Notes in Computer Science*, *3176*, 63–71. https://doi.org/10.1007/978-3-540-28650-9_4

Rastetter, E. B., King, A. W., Cosby, B. J., Hornberger, G. M., Oneill, R. V., & Hobbie, J. E. (1992). Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. *Ecological Applications*, *2*(1), 55–70. https://doi.org/10.2307/1941889

Rawles, W., & Brakensiek, D. (1982). Estimating soil water retention from soil properties. *Journal of the Irrigation and Drainage Division*, *108*(2), 166–171.

Rawls, W. J., & Brakensiek, D. L. (1985). Prediction of soil water properties for hydrologic modeling. In *Watershed management in the eighties* (pp. 293–299). Reston, VA: American Society of Civil Engineers.

Rawls, W. J., & Pachepsky, Y. A. (2002a). Soil consistence and structure as predictors of water retention. *Soil Science Society of America Journal*, *66*(4), 1115–1126. https://doi.org/10.2136/sssaj2002.1115

Rawls, W. J., & Pachepsky, Y. A. (2002b). Using field topographic descriptors to estimate soil water retention. *Soil Science*, *167*(7), 423–435. https://doi.org/10.1097/00010694-200207000-00001

Rawls, W. J., Brakensiek, D. L., & Soni, B. (1983). Agricultural management effects on soil water processes. Part I: Soil water retention and Green and Ampt infiltration parameters. *Transactions of ASAE*, *26*(6), 1747–1752. https://doi.org/10.13031/2013.33837

Rewald, B., Ephrath, J. E., & Rachmilevitch, S. (2011). A root is a root is a root? Water uptake rates of *Citrus* root orders. *Plant, Cell & Environment*, *34*(1), 33–42. https://doi.org/10.1111/j.1365-3040.2010.02223.x

Rey, A., Oyonarte, C., Morán-López, T., Raimundo, J., & Pegoraro, E. (2017). Changes in soil moisture predict soil carbon losses upon rewetting in a perennial semiarid steppe in SE Spain. *Geoderma*, *287*, 135–146. https://doi.org/10.1016/j.geoderma.2016.06.025

Richards, L. A. (1931). Capillary conduction of liquids through porous mediums. *Physics*, *1*(5), 318–333. https://doi.org/10.1063/1.1745010

Ritsema, C. J., & Dekker, L. W. (1996). Water repellency and its role in forming preferred flow paths in soils. *Australian Journal of Soil Research*, *34*(4), 475–487. https://doi.org/10.1071/SR9960475

Robinson, D. A., Jones, S. B., Lebron, I., Reinsch, S., Domínguez, M. T., Smith, A. R., … Emmett, B. A. (2016). Experimental evidence for drought induced alternative stable states of soil moisture. *Scientific Reports*, *6*(1), 20018. https://doi.org/10.1038/srep20018

Rodell, M., Houser, P. R., Jambor, U. E. A., Gottschalck, J., Mitchell, K., Meng, C. J., … Bosilovich, M. (2004). The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society*, *85*(3), 381–394. https://doi.org/10.1175/BAMS-85-3-381

Rodriguez, R. J., Henson, J., Van Volkenburgh, E., & Hoy, M. (2008). Stress tolerance in plants via habitat-adapted symbiosis. *ISME Journal*, *2*(4), 404–416. https://doi.org/10.1038/ismej.2007.106

Rodriguez-Iturbe, I., & Porporato, A. (2004). *Ecohydrology of water controlled ecosystems: Soil moisture and plant dynamics*. New York: Cambridge University Press.

Rodríguez-Lado, L., Rial, M., Taboada, T., & Cortizas, A. M. (2015). A pedotransfer function to map soil bulk density from limited data. *Procedia Environmental Sciences*, *27*, 45–48. https://doi.org/10.1016/j.proenv.2015.07.112

Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., and Tompkins, A. (2003). The atmospheric general circulation model ECHAM 5. PART I: Model description, MPI Report 349.

Rogers, A., Medlyn, B. E., Dukes, J. S., Bonan, G., von Caemmerer, S., Dietze, M. C., … Zaehle, S. (2016). A roadmap for improving the representation of photosynthesis in Earth system models. *The New Phytologist*, *213*(1), 22–42. https://doi.org/10.1111/nph.14283

Romano, N. (2004). Spatial structure of PTF estimates. In Y. Pachepsky & W. J. Rawls (Eds.), *Development of pedotransfer functions in soil hydrology* (1st ed., Vol. 30, pp. 295–315). Springer Verlag.

Romano, N., & Palladino, M. (2002). Prediction of soil water retention using soil physical data and terrain attributes. *Journal of Hydrology*, *265*(1-4), 56–75. https://doi.org/10.1016/S0022-1694(02)00094-X

Roose, T., & Fowler, A. C. (2004). A model for water uptake by plant roots. *Journal of Theoretical Biology*, *228*, 151–171. https://doi.org/10.1016/j.jtbi.2003.12.012

Roth, K. (2008). Scaling of water flow through porous media and soils. *European Journal of Soil Science*, *59*(1), 125–130. https://doi.org/10.1111/j.1365-2389.2007.00986.x

Rubin, Y., Chen, X., Murakami, H., & Hahn, M. S. (2010). A Bayesian approach for inverse modeling, data assimilation, and conditional simulation of spatial random fields. *Water Resources Research*, *46*(10), W10523. https://doi.org/10.1029/2009WR008799

Rudiyanto, S., Minasny, B., & Setiawan, B. I. (2016). Further results on comparison of methods for quantifying soil carbon in tropical peats. *Geoderma*, *269*, 108–111. https://doi.org/10.1016/j.geoderma.2016.01.038

Sakurai, G., Jomura, M., Yonemura, S., Iizumi, T., Shirato, Y., & Yokozawa, M. (2012). Inversely estimating temperature sensitivity of soil carbon decomposition by assimilating a turnover model and long-term field data. *Soil Biology and Biochemistry*, *46*, 191–199. https://doi.org/10.1016/j.soilbio.2011.11.005

Samaniego, L., Bárdossy, A., & Kumar, R. (2010). Streamflow prediction in ungauged catchments using copula-based dissimilarity measures. *Water Resources Research*, *46*, W02506. https://doi.org/10.1029/2008WR007695

Samaniego, L., Kumar, R., & Attinger, S. (2010). Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resources Research*, *46*, W05523. https://doi.org/10.1029/2008WR007327

Sanchez, P. A., Ahamed, S., Carré, F., Hartemink, A. E., Hempel, J., Huising, J., … Zhang, G. L. (2009). Digital soil map of the world. *Science*, *325*(5941), 680–681. https://doi.org/10.1126/science.1175084

Sánchez-Vila, X., Carrera, J., & Girardi, J. P. (1996). Scale effects in transmissivity. *Journal of Hydrology*, *183*(1-2), 1–22. https://doi.org/10.1016/S0022-1694(96)80031-X

Sato, H., Ito, A., Ito, A., Ise, T., & Kato, E. (2015). Current status and future of land surface models. *Soil Science & Plant Nutrition*, *61*(1), 34–47. https://doi.org/10.1080/00380768.2014.917593

Saxton, K. E., & Rawls, W. J. (2006). Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Science Society of America Journal*, *70*(5), 1569–1578. https://doi.org/10.2136/sssaj2005.0117

Saxton, K. E., Rawls, W., Romberger, J. S., & Papendick, R. I. (1986). Estimating generalized soil-water characteristics from texture. *Soil Science Society of America Journal*, *50*(4), 1031–1036. https://doi.org/10.2136/sssaj1986.03615995005000040039x

Schaap, M. G. (2004). Accuracy and uncertainty in PTF predictions. In Y. A. Pachepsky & W. J. Rawls (Eds.), *Developments in soil science* (Vol. 30, pp. 33–43). Amsterdam: Elsevier.

Schaap, M. G., & Bouten, W. (1996). Modeling water retention curves of sandy soils using neural networks. *Water Resources Research*, *32*(10), 3033–3040. https://doi.org/10.1029/96WR02278

Schaap, M. G., & Leij, F. J. (1998a). Database-related accuracy and uncertainty of pedotransfer functions. *Soil Science*, *163*(10), 765–779. https://doi.org/10.1097/00010694-199810000-00001

Schaap, M. G., & Leij, F. J. (1998b). Using neural networks to predict soil water retention and soil hydraulic conductivity. *Soil and Tillage Research*, *47*(1-2), 37–42. https://doi.org/10.1016/S0167-1987(98)00070-1

Schaap, M. G., & van Genuchten, M. T. (2006). A modified Mualem–van Genuchten formulation for improved description of the hydraulic conductivity near saturation. *Vadose Zone Journal*, *5*(1), 27. https://doi.org/10.2136/vzj2005.0005

Schaap, M. G., Leij, F. J., & van Genuchten, M. T. (1998). Neural network analysis for hierarchical prediction of soil hydraulic properties. *Soil Science Society of America Journal*, *62*(4), 847–855. https://doi.org/10.2136/sssaj1998.03615995006200040001x

Schaap, M. G., Leij, F. J., & van Genuchten, M. T. (2001). ROSETTA: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *Journal of Hydrology*, *251*(3-4), 163–176. https://doi.org/10.1016/S0022-1694(01)00466-8

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*(4), 1455–1463. https://doi.org/10.2136/vzj2004.1455

Scheidegger, A. E. (1954). Statistical hydrodynamics in porous media. *Journal of Applied Physics*, *25*(8), 994–1001. https://doi.org/10.1063/1.1721815

Schenk, H. J. (2008). The shallowest possible water extraction profile: A null model for global root distributions. *Vadose Zone Journal*, *7*(3), 1119–1124. https://doi.org/10.2136/vzj2007.0119

Schenk, H. J., & Jackson, R. B. (2002). The global biogeography of roots. *Ecological Monographs*, *72*(3), 311–328. https://doi.org/10.2307/3100092

Schenk, H. J., & Jackson, R. B. (2005). Mapping the global distribution of deep roots in relation to climate and soil characteristics. *Geoderma*, *126*(1-2), 129–140. https://doi.org/10.1016/j.geoderma.2004.11.018

Schimel, D., Melillo, J., Tian, H., McGuire, A. D., Kicklighter, D., Kittel, T., & Rizzo, B. (2000). Contribution of increasing CO_2 and climate to carbon storage byecosystems in the United States. *Science*, *287*(5460), 2004–2006. https://doi.org/10.1126/science.287.5460.2004

Schrumpf, M., Schulze, E. D., Kaiser, K., & Schumacher, J. (2011). How accurately can soil organic carbon stocks and stock changes be quantified by soil inventories? *Biogeosciences*, *8*(5), 1193–1212. https://doi.org/10.5194/bg-8-1193-2011

Schulze-Makuch, D., Carlson, D. A., Cherkauer, D. S., & Malik, P. (1999). Scale dependency of hydraulic conductivity in heterogeneous media. *Ground Water*, *37*(6), 904–919. https://doi.org/10.1111/j.1745-6584.1999.tb01190.x

Schwartz, N., Carminati, A., & Javaux, M. (2016). The impact of mucilage on root water uptake—A numerical study. *Water Resources Research*, *52*(1), 264–277. https://doi.org/10.1002/2015WR018150

Schymanski, S. J., Sivapalan, M., Roderick, M. L., Beringer, J., & Hutley, L. B. (2008). An optimality-based model of the coupled soil moisture and root dynamics. *Hydrology and Earth System Sciences*, *12*(3), 913–932. https://doi.org/10.5194/hess-12-913-2008

Segal, M. R. (2004). Machine learning benchmarks and random forest regression. Center for Bioinformatics and Molecular Biostatistics.

Sellers, P. J., Dickinson, R. E., Randall, D. A., Betts, A. K., Hall, F. G., Berry, J. A., & Henderson-Sellers, A. (1997). Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science*, *275*(5299), 502–509. https://doi.org/10.1126/science.275.5299.502

Seneviratne, S. I., Koster, R. D., Guo, Z., Dirmeyer, P. A., Kowalczyk, E., Lawrence, D., … Verseghy, D. (2006). Soil moisture memory in AGCM simulations: Analysis of Global Land-Atmosphere Coupling Experiment (GLACE) data. *Journal of Hydrometeorology*, *7*(5), 1090–1112. https://doi.org/10.1175/JHM533.1

Sequeira, C. H., Wills, S. A., Seybold, C. A., & West, L. T. (2014). Predicting soil bulk density for incomplete databases. *Geoderma*, *213*, 64–73. https://doi.org/10.1016/j.geoderma.2013.07.013

Serna-Chavez, H. M., Fierer, N., & van Bodegom, P. M. (2013). Global drivers and patterns of microbial abundance in soil. *Global Ecology and Biogeography*, *22*(10), 1162–1172. https://doi.org/10.1111/geb.12070

Sharma, S. K., Mohanty, B. P., & Zhu, J. (2006). Including topography and vegetation attributes for developing pedotransfer functions. *Soil Science Society of America Journal*, *70*(5), 1430–1440. https://doi.org/10.2136/sssaj2005.0087

Shaw, J. N., West, L. T., Radcliffe, D. E., & Bosch, D. D. (2000). Preferential flow and pedotransfer functions for transport properties in sandy kandiudults. *Soil Science Society of America Journal*, *64*(2), 670–678. https://doi.org/10.2136/sssaj2000.642670x

Shen, C., Niu, J., & Fang, K. (2014). Quantifying the effects of data integration algorithms on the outcomes of a subsurface–land surface processes model. *Environmental Modelling and Software*, *59*, 146–161. https://doi.org/10.1016/j.envsoft.2014.05.006

Sherman, C., Sternberg, M., & Steinberger, Y. (2012). Effects of climate change on soil respiration and carbon processing in Mediterranean and semi-arid regions: An experimental approach. *European Journal of Soil Biology*, *52*, 48–58. https://doi.org/10.1016/j.ejsobi.2012.06.001

Shiri, J., Keshavarzi, A., Kisi, O., Karimi, S., & Iturraran-Viveros, U. (2017). Modeling soil bulk density through a complete data scanning procedure: Heuristic alternatives. *Journal of Hydrology*, *549*, 592–602. https://doi.org/10.1016/j.jhydrol.2017.04.035

Simmons, C. S. (1982). A stochastic-convective transport representation of dispersion in one-dimensional porous-media systems. *Water Resources Research*, *18*(4), 1193–1214. https://doi.org/10.1029/WR018i004p01193

Šimůnek, J., Jacques, D., Twarakavi, N. K. C., & van Genuchten, M. T. (2009). Selected HYDRUS modules for modeling subsurface flow and contaminant transport as influenced by biological processes at various scales. *Biologia*, *64*(3), 465–469.

Skalová, J., Čistý, M., & Bezák, J. (2011). Comparison of three regression models for determining water retention curves. *Journal of Hydrological Hydromechanics*, *59*(4), 275–284.

Smettem, K., Pracilio, G., Oliver, Y., & Harper, R. (2004). Data availability and scale in hydrologic applications. In Y. Pachepsky & W. J. Rawls (Eds.), *Development of pedotransfer functions in soil hydrology* (pp. 253–271). Elsevier. https://doi.org/10.1016/S0166-2481(04)30015-2

Solomatine, D. P., & Dulal, K. N. (2003). Model trees as an alternative to neural networks in rainfall—Runoff modelling. *Hydrological Sciences Journal*, *48*(3), 399–411. https://doi.org/10.1623/hysj.48.3.399.45291

Somma, F., Clausnitzer, V., & Hopmans, J. W. (1998). An algorithm for three-dimensional simultaneous modeling of root growth, transient soil water flow, and transport and uptake, V.2.1. Land, Air and Water Resources Paper No. 100034, University of California, Davis.

Sourbeer, J. J., & Loheide, S. P. II (2015). Obstacles to long-term soil moisture monitoring with heated distributed temperature sensing. *Hydrological Processes*, *30*, 1017–1035.

Stacke, T., & Hagemann, S. (2012). Development and evaluation of a global dynamical wetlands extent scheme. *Hydrology and Earth System Sciences*, *16*(8), 2915–2933. https://doi.org/10.5194/hess-16-2915-2012

Stamati, F. E., Nikolaidis, N. P., Banwart, S., & Blum, W. E. H. (2013). A coupled carbon, aggregation, and structure turnover (CAST) model for topsoils. *Geoderma*, *211-212*(1), 51–64. https://doi.org/10.1016/j.geoderma.2013.06.014

Staver, A. C., & Hansen, M. C. (2015). Analysis of stable states in global savannas: Is the CART pulling the horse?–a comment. *Global Ecology and Biogeography*, *24*(8), 985–987. https://doi.org/10.1111/geb.12285

Stolte, J., Wesseling, J. G., & Wösten, J. H. M. (1996). Pedotransfer functions for hydraulic and thermal properties of soil and the tool HERCULES, Wageningen, SC-DLO, Report 126 (41 pp.).

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., & Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics*, *9*(1), 307. https://doi.org/10.1186/1471-2105-9-307

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application and characteristics of classification and regression trees, bagging and random forests. *Psychological Methods*, *14*(4), 323–348. https://doi.org/10.1037/a0016973

Subin, Z. M., Koven, C. D., Riley, W. J., Torn, M. S., Lawrence, D. M., & Swenson, S. C. (2013). Effects of soil moisture on the responses of soil temperatures to climate change in cold regions. *Journal of Climate*, *26*(10), 3139–3158. https://doi.org/10.1175/JCLI-D-12-00305.1

Sulman, B. N., Phillips, R. P., Oishi, A. C., Shevliakova, E. & Pacala, S. W. (2014). Microbe-driven turnover offsets mineral-mediated storage of soil carbon under elevated CO_2. *Nature Climate Change*, *4*, 1099.

Sun, S., & Xue, Y. (2001). Implementing a new snow scheme in Simplified Simple Biosphere Model (SSiB). *Advances in Atmospheric Sciences*, *18*, 335–354.

Suwardji, P., & Eberbach, P. L. (1998). Seasonal changes in soil physical properties of an Oxic Paleustalf (Red Kandosol) after 16 years of direct drilling and conventional cultivation. *Soil and Tillage Research*, *49*(1-2), 65–77. https://doi.org/10.1016/S0167-1987(98)00149-4

Tamari, S., Wösten, J. H. M., & Ruiz-Suarez, J. C. (1996). Testing an artificial neural network for predicting soil hydraulic conductivity. *Soil Science Society of America Journal*, *60*(6), 1732–1741. https://doi.org/10.2136/sssaj1996.03615995006000060018x

Tang, J., Riley, W. J., & Niu, J. (2015). Incorporating root hydraulic redistribution in CLM4.5: Effects on predicted site and global evapotranspiration, soil moisture and water storage. *Journal of Advances in Modeling Earth Systems*, *7*(4), 1828–1848. https://doi.org/10.1002/2015MS000484

Tarnawski, V. R., Momose, T., & Leong, W. H. (2009). Assessing the impact of quartz content on the prediction of soil thermal conductivity. *Géotechnique*, *59*(4), 331–338. https://doi.org/10.1680/geot.2009.59.4.331

Tarnawski, V. R., Momose, T., & Leong, W. H. (2011). Thermal conductivity of standard sands II. Saturated conditions. *International Journal of Thermophysics*, *32*(5), 984–1005. https://doi.org/10.1007/s10765-011-0975-1

Tartakovsky, A. M., Scheibe, T. D., & Meakin, P. (2009). Pore-scale model for reactive transport and biomass growth. *Journal of Porous Media*, *12*(5), 417–434. https://doi.org/10.1615/JPorMedia.v12.i5.30

Tempel, P., Batjes, N. H., & van Engelen, V. W. P. (1996). IGBP-DIS soil data set for pedotransfer function development. Int. Soil Reference and Inf. Centre, Wageningen, the Netherlands.

Thomas, R. Q., Zaehle, S., Templer, P. H., & Goodale, C. L. (2013). Global patterns of nitrogen limitation: Confronting two global biogeochemical models with observations. *Global Change Biology*, *19*(10), 2986–2998. https://doi.org/10.1111/gcb.12281

Thomas, R. Q., Brookshire, E. N. J., & Gerber, S. (2015). Nitrogen limitation on land: How can it occur in Earth system models? *Global Change Biology*, *21*(5), 1777–1793. https://doi.org/10.1111/gcb.12813

Tian, H., Melillo, J. M., Kicklighter, D. W., McGuire, A. D., & Helfrich, J. V. K. I. (1999). The sensitivity of terrestrial carbon storage to historical climate variability and atmospheric CO_2 in the United States. *Tellus*, *51*(2), 414–452. https://doi.org/10.3402/tellusb.v51i2.16318

Tietje, O., & Hennings, V. (1996). Accuracy of the saturated hydraulic conductivity prediction by pedo-transfer functions compared to the variability within FAO textural classes. *Geoderma*, *69*(1-2), 71–84. https://doi.org/10.1016/0016-7061(95)00050-X

Tietje, O., & Tapkenhinrichs, M. (1993). Evaluation of pedo-transfer functions. *Soil Science Society of America Journal*, *57*(4), 1088–1095. https://doi.org/10.2136/sssaj1993.03615995005700040035x

Timilsina, N., Escobedo, F. J., Cropper, W. P. Jr., Abd-Elrahman, A., Brandeis, T. J., Delphin, S., & Lambert, S. (2013). A framework for identifying carbon hotspots and forest management drivers. *Journal of Environmental Management*, *114*, 293–302. https://doi.org/10.1016/j.jenvman.2012.10.020

Timlin, D. J., Ahuja, L. R., Pachepsky, Y. A., Williams, R. D., Gimenez, D., & Rawls, W. J. (1999). Use of Brooks-Corey parameters to improve estimates of saturated conductivity from effective porosity. *Soil Science Society of America Journal*, *63*(5), 1086–1092. https://doi.org/10.2136/sssaj1999.6351086x

Tóth, B., Makó, A., Guadagnini, A., & Tóth, G. (2012). Water retention of salt affected soils: Quantitative estimation using soil survey information. *Arid Land Research and Management*, *26*(2), 103–121. https://doi.org/10.1080/15324982.2012.657025

Tóth, B., Makó, A., & Tóth, G. (2014). Role of soil properties in water retention characteristics of main Hungarian soil types. *Journal of Central European Agriculture*, *15*(2), 137–153. https://doi.org/10.5513/JCEA01/15.2.1465

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*(1), 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*(14), 2662–2666. https://doi.org/10.1002/hyp.11203

Tranter, G., McBratney, A. B., & Minasny, B. (2009). Using distance metrics to determine the appropriate domain of pedotransfer function predictions. *Geoderma*, *149*(3-4), 421–425. https://doi.org/10.1016/j.geoderma.2009.01.006

Tranter, G., Minasny, B., & McBratney, A. B. (2010). Estimating pedotransfer function prediction limits using fuzzy-means with extragrades. *Soil Science Society of America Journal*, *74*(6), 1967–1975. https://doi.org/10.2136/sssaj2009.0106

Tuller, M., & Or, D. (2001). Hydraulic conductivity of variably saturated porous media: Film and corner flow in angular porous media. *Water Resources Research*, *37*(5), 1257–1276. https://doi.org/10.1029/2000WR900328

Twarakavi, N. K. C., Šimůnek, J., & Schaap, M. G. (2009). Development of pedotransfer functions for estimation of soil hydraulic parameters using support vector machines. *Soil Science Society of America Journal*, *73*(5), 1443–1452. https://doi.org/10.2136/sssaj2008.0021

Twarakavi, N. K. C., Simunek, J., & Schaap, M. G. (2010). Can texture-based classification optimally classify soils with respect to soil hydraulics? *Water Resources Research*, *46*, W01501.

van Wijk, W. R., & de Vries, D. A. (1963). Periodic temperature variations in a homogenous soil. In van Wijk W. R. (Ed.), *Physics of Soil Environment* (pp. 102–143). Amsterdam: North-Holland Publishing Company.

van Wijk, M. T., & Bouten, W. (2001). Towards understanding tree root profiles: Simulating hydrologically optimal strategies for root distribution. *Hydrology and Earth System Sciences*, *5*(4), 629–644. https://doi.org/10.5194/hess-5-629-2001

van Bemmelen, J. M. (1890). Ueber die Bestimmung des Wassers, des Humus, des Schwefels, der in den kolloidalen Silikaten gebundenen Kieselsäure, und des Mangans, im Ackerboden. In *Die Landwirtschaftlichen Versuchs-Stationen* (Vol. 37, p. 277). Chemnitz.

Vanderborght, J., & Vereecken, H. (2007a). One-dimensional modeling of transport in soils with depth-dependent dispersion, sorption and decay. *Vadose Zone Journal*, *6*(1), 140–148. https://doi.org/10.2136/vzj2006.0103

Vanderborght, J., & Vereecken, H. (2007b). Review of dispersivities for transport modeling in soils. *Vadose Zone Journal*, *6*(1), 29–52. https://doi.org/10.2136/vzj2006.0096

Vanderborght, J., Vanclooster, M., Timmerman, A., Seuntjens, P., Mallants, D., Kim, D. J., … Deckers, J. (2001). Overview of inert tracer experiments in key Belgian soil types: Relation between transport and soil morphological and hydraulic properties. *Water Resources Research*, *37*(12), 2873–2888. https://doi.org/10.1029/2000WR000110

van Genuchten, M. T. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, *44*(5), 892–898. https://doi.org/10.2136/sssaj1980.03615995004400050002x

van Tol, J. J., Le Roux, P. A. L., Lorentz, S. A., & Nensley, M. (2013). Hydropedological classification of South African hillslopes. *Vadose Zone Journal*, *12*(4). https://doi.org/10.2136/vzj2013.01.0007

Vapnik, V. (2013). *The nature of statistical learning theory*. New York: Springer Science and Business Media.

Vapnik, V. N., & Vapnik, V. (1998). *Statistical learning theory*. New York: Wiley.

Vasilyeva, N. A., Ingtem, J. G., & Silaev, D. A. (2016). Nonlinear dynamical model of microorganism growth in soil. *Computational Mathematics and Modeling*, *27*(2), 172–180. https://doi.org/10.1007/s10598-016-9312-7

Veihmeyer, F. J., & Hendrickson, A. H. (1927). The relation of soil moisture to cultivation and plant growth. Proc. 1th intern. *Congress of Soil Science*, *3*, 498–513.

Vereecken, H. (2002). Comment on the paper, "Evaluation of pedo-transfer functions for unsaturated soil hydraulic conductivity using an independent data set". *Geoderma*, *108*(1-2), 145–147. https://doi.org/10.1016/S0016-7061(02)00127-1

Vereecken, H., Maes, J., Feyen, J., & Darius, P. (1989). Estimating the soil moisture retention characteristic from texture, bulk density, and carbon content. *Soil Science*, *148*(6), 389–403. https://doi.org/10.1097/00010694-198912000-00001

Vereecken, H., Maes, J., & Feyen, J. (1990). Estimating unsaturated hydraulic conductivity from easily measured soil properties. *Soil Science*, *149*(1), 1–12. https://doi.org/10.1097/00010694-199001000-00001

Vereecken, H., Diels, J., Van Orshoven, J., Feyen, J., & Bouma, J. (1992). Functional evaluation of pedotransfer functions for the estimation of soil hydraulic properties. *Soil Science Society of America Journal*, *56*(5), 1371–1378. https://doi.org/10.2136/sssaj1992.03615995005600050007x

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*(4), 795. https://doi.org/10.2136/vzj2010.0045

Vereecken, H., Vanderborght, J., Kasteel, R., Spiteller, M., Schaffer, A., & Close, M. (2011). Do lab-derived distribution coefficient values of pesticides match distribution coefficient values determined from column and field-scale experiments? A critical analysis of relevant literature. *Journal of Environmental Quality*, *40*(3), 879–898. https://doi.org/10.2134/jeq2010.0404

Vereecken, H., Schnepf, A., Hopmans, J. W., Javaux, M., Or, D., Roose, T., … Young, I. M. (2016). Modeling soil processes: Review, key challenges, and new perspectives. *Vadose Zone Journal*, *15*(5). https://doi.org/10.2136/vzj2015.09.0131

Vitharana, U. W. A., Mishra, U., Jastrow, J. D., Matamala, R., & Fan, Z. (2017). Observational needs for estimating Alaskan soil carbon stocks under current and future climate. *Journal of Geophysical Research: Biogeosciences*, *122*, 415–429. https://doi.org/10.1002/2016JG003421

Vogeler, I., Cichota, R., Snow, V. O., Dutton, T., & Daly, B. (2011). Pedotransfer functions for estimating ammonium adsorption in soils. *Soil Science Society of America Journal*, *75*(1), 324–331. https://doi.org/10.2136/sssaj2010.0192

von Arx, G., Archer, S. R., & Hughes, M. K. (2012). Long-term functional plasticity in plant hydraulic architecture in response to supplemental moisture. *Annals of Botany*, *109*(6), 1091–1100. https://doi.org/10.1093/aob/mcs030

von Gotz, N., & Richter, O. (1999). Simulation of herbicide degradation in different soils by use of pedo-transfer functions (PTF) and non-linear kinetics. *Chemosphere*, *38*(6), 1401–1407. https://doi.org/10.1016/s0045-6535(98)00542-6

Walko, R. L., Band, L. E., Baron, J., Kittel, T. G. F., Lammers, R., Lee, T. J., … Vidale, P. L. (2000). Coupled atmosphere-biophysics-hydrology models for environmental modeling. *Journal of Applied Meteorology*, *39*(6), 931–944. https://doi.org/10.1175/1520-0450(2000)039%3C0931:CABHMF%3E2.0.CO;2

Wang, Y. P., Kowalczyk, E., Leuning, R., Abramowitz, G., Raupach, M. R., Pak, B., … Luhar, A. (2011). Diagnosing errors in a land surface model (CABLE) in the time and frequency domains. *Journal of Geophysical Research*, *116*, G01034. https://doi.org/10.1029/2010JG001385

Wang, G., Luo, Z., Han, P., Chen, H., & Xu, J. (2016). Critical carbon input to maintain current soil organic carbon stocks in global wheat systems. *Scientific Reports*, *6*(1), 19327. https://doi.org/10.1038/srep19327

Weihermüller, L., Graf, A., Herbst, M., & Vereeecken, H. (2013). Simple pedotransfer functions to initialize reactive carbon pools of the RothC model. *European Journal of Soil Science*, *64*(5), 567–575. https://doi.org/10.1111/ejss.12036

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*(1). https://doi.org/10.2136/vzj2008.0062er

Weynants, M., Vereecken, H., & Javaux, M. (2009). Revisiting Vereecken pedotransfer functions: Introducing a closed-form hydraulic model. *Vadose Zone Journal*, *8*(1), 86. https://doi.org/10.2136/vzj2008.0062

Weynants, M., Montanarella, L., Toth, G., Strauss, P., Feichtinger, F., Cornelis, W., … Henk, W. (2013). *European HYdropedological Data Inventory (EU-HYDI)*. Luxembourg: EUR – Scientific and Technical Research Series. Publications Office of the European Union.

Whitfield, C. J., & Reid, C. (2013). Predicting surface area of coarse-textured soils: Implications for weathering rates. *Canadian Journal of Soil Science*, *93*(5), 621–630. https://doi.org/10.4141/cjss2012-009

Wieder, W. R., Bonan, G. B., & Allison, S. D. (2013). Global soil carbon projections are improved by modelling microbial processes. *Nature Climate Change*, *3*, 909.

Williams, R. D., & Ahuja, L. R. (2003). Scaling and estimating the soil water characteristic using a one-parameter model. In Y. Pachepsky, D. E. Radcliffe, & H. M. Selim (Eds.), *Scaling Methods in Soil Physics* (pp. 35–48). Boca Raton, FL: CRC Press.

Williams, J., Ross, P. J., & Bristow, K. L. (1992). Prediction of the Campbell water retention function from texture, structure and organic matter. In M. T. van Genuchten, F. J. Leij, & L. J. Lund (Eds.), *Proceedings of the international workshop on indirect methods for estimating the hydraulic properties of unsaturated soils* (pp. 427–441). Riverside, CA: University of California.

Wösten, J. H. M., Bannink, M. H., de Gruijter, J. J., & Bouma, J. (1986). A procedure to identify different groups of hydraulic conductivity and moisture retention curves for soil horizons. *Journal of Hydrology*, *86*(133), 145.

Wösten, J. H. M., Schuren, C. H. E. J., Bouma, J., & Stein, A. (1990). Functional sensitivity analysis of four methods to generate soil hydraulic functions. *Soil Science Society of America Journal*, *54*, 827–832.

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*(3-4), 169–185. https://doi.org/10.1016/S0016-7061(98)00132-3

Wösten, J. H. M., Pachepsky, Y., & Rawls, W. J. (2001). Pedotransfer functions: Bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*, *251*(3-4), 123–150. https://doi.org/10.1016/S0022-1694(01)00464-4

Xevi, E., Christiaens, K., Espino, A., Sewnandan, W., Mallants, D., Sørensen, H., & Feyen, J. (1997). Calibration, validation and sensitivity analysis of the MIKE-SHE model using the Neuenkirchen catchment as case study. *Water Resources Management*, *11*(3), 219–242. https://doi.org/10.1023/A:1007977521604

Xiangsheng, Y., Guosheng, L., & Yanyu, Y. (2016). Pedotransfer functions for estimating soil bulk density: A case study in the Three-River Headwater region of Qinghai Province, China. *Pedosphere*, *26*(3), 362–373.

Xu, M., & Qi, Y. (2001). Spatial and seasonal variations of Q10 determined by soil respiration measurements at a Sierra Nevadan forest. *Global Biogeochemical Cycles*, *15*(3), 687–696. https://doi.org/10.1029/2000GB001365

Xu, X., Thornton, P. E., & Post, W. M. (2013). A global analysis of soil microbial biomass carbon, nitrogen and phosphorus in terrestrial ecosystems. *Global Ecology and Biogeography*, *22*(6), 737–749. https://doi.org/10.1111/geb.12029

Xu, L., He, N. P., Yu, G. R., Wen, D., Gao, Y., & He, H. L. (2015). Differences in pedotransfer functions of bulk density lead to high uncertainty in soil organic carbon estimation at regional scales: Evidence from Chinese terrestrial ecosystems. *Journal of Geophysical Research: Biogeosciences*, *120*, 1567–1575.

Xue, Y., Sellers, P. J., Kinter, J. L. III, & Shukla, J. (1991). A simplified biosphere model for global climate studies. *Journal of Climate*, *4*(3), 345–364. https://doi.org/10.1175/1520-0442(1991)004%3C0345:ASBMFG%3E2.0.CO;2

Xu-Ri, X., & Prentice, I. C. (2008). Terrestrial nitrogen cycle simulation with a dynamic global vegetation model. *Global Change Biology*, *14*(8), 1745–1764. https://doi.org/10.1111/j.1365-2486.2008.01625.x

Yang, Y. T., Donohue, R. J., & McVicar, T. R. (2016). Global estimation of effective plant rooting depth: Implications for hydrological modeling. *Water Resources Research*, *52*(10), 8260–8276. https://doi.org/10.1002/2016wr019392

Ye, M., Khaleel, R., Schaap, M. G., & Zhu, J. (2007). Simulation of field injection experiments in heterogeneous unsaturated media using cokriging and artificial neural network. *Water Resources Research*, *43*(7), W07413. https://doi.org/10.1029/2006WR005030

Yi, X., Li, G., & Yin, Y. (2013). Comparison of three methods to develop pedotransfer functions for the saturated water content and field water capacity in permafrost region. *Cold Regions Science and Technology*, *88*, 10–16. https://doi.org/10.1016/j.coldregions.2012.12.005

Young, M. D. B., Gowing, J. W., Wyseure, G. C. L., & Hatibu, N. (2002). Parched–Thirst: Development and validation of a process-based model of rainwater harvesting. *Agricultural Water Management*, *55*(2), 121–140. https://doi.org/10.1016/S0378-3774(01)00187-1

Zacharias, S., & Bohne, K. (2008). Attempt of a flux-based evaluation of field capacity. *Journal of Plant Nutrition and Soil Science*, *171*(3), 399–408. https://doi.org/10.1002/jpln.200625168

Zacharias, S., & Wessolek, G. (2007). Excluding organic matter content from pedotransfer predictors of soil water retention. *Soil Science Society of America Journal*, *71*(1), 43–50. https://doi.org/10.2136/sssaj2006.0098

Zaehle, S., & Dalmonech, D. (2011). Carbon-nitrogen interactions on land at global scales: Current understanding in modelling climate biosphere feedbacks. *Current Opinion in Environment Sustainability*, *3*(5), 311–320. https://doi.org/10.1016/j.cosust.2011.08.008

Zaehle, S., & Friend, A. (2010). Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates. *Global Biochemical Cycles*, *24*, GB1005. https://doi.org/10.1029/2009GB003521

Zaller, J. G., Heigl, F., Grabmaier, A., Lichtenegger, C., Piller, K., Allabashi, R., … Drapela, T. (2011). Earthworm-Mycorrhiza interactions can affect the diversity, structure and functioning of establishing model grassland communities. *PLoS One*, *6*(12), e29293. https://doi.org/10.1371/journal.pone.0029293

Zhan, X., Xue, Y., & Collatz, G. J. (2003). An analytical approach for estimating CO_2 and heat fluxes over the Amazonian region. *Ecological Modeling*, *162*(1-2), 97–117. https://doi.org/10.1016/S0304-3800(02)00405-2

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

Zhao, C., Miao, Y., Yu, C., Zhu, L., Wang, F., Jiang, L., … Wan, S. (2016). Soil microbial community composition and respiration along an experimental precipitation gradient in a semiarid steppe. *Scientific Reports*, *6*(1), 24317. https://doi.org/10.1038/srep24317

Zhou, T., Shi, P., Hui, D., & Luo, Y. (2009). Global pattern of temperature sensitivity of soil heterotrophic respiration (Q10) and its implications for carbon-climate feedback. *Journal of Geophysical Research*, *114*, G02016. https://doi.org/10.1029/2008JG000850

Zhu, G., & Blumberg, D. G. (2002). Classification using ASTER data and SVM algorithms: The case study of beer Sheva, Israel. *Remote Sensing of Environment*, *80*(2), 233–240. https://doi.org/10.1016/S0034-4257(01)00305-4

Zimmermann, M., Leifeld, J., Schmidt, M. W., Smith, P., & Fuhrer, J. (2007). Measured soil organic matter fractions can be related to pools in the RothC model. *European Journal of Soil Science*, *58*(3), 658–667. https://doi.org/10.1111/j.1365-2389.2006.00855.x