Development of Hierarchical Ensemble Model and Estimates of Soil Water Retention With Global Coverage

Yonggen Zhang1,2, Marcel G. Schaap3, and Zhongwang Wei3,4

1Institute of Surface-Earth System Science, School of Earth System Science, Tianjin University, Tianjin, China, 2Department of Soil, Water and Environmental Science, University of Arizona, Tucson, AZ, USA, 3Guangdong Province Key Laboratory for Climate Change and Natural Disaster Studies, School of Atmospheric Sciences, Sun Yat-sen University, Guangzhou, China, 4Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, China

Abstract Correct quantification of mass and energy exchange processes between land surface and atmosphere requires an accurate description of unsaturated soil hydraulic properties. Soil pedotransfer functions (PTFs) have been widely used to predict soil hydraulic parameters. Here, 13 PTFs were grouped according to input data requirements and evaluated against a well-documented database (National Cooperative Soil Survey Characterization [NCSS]) covering the continental United States (87.7% of data) and other regions of the globe (12.3% of data). Weighted ensembles were shown to have improved performance over individual PTFs in terms of evaluation criteria. Validation of moisture content estimated from the ensemble models against observations showed promising results. Global maps of soil water retention data from the ensemble models as well as their uncertainty were provided. Our full 13-member ensemble model provides more accurate estimates than PTFs that are currently being used in Earth system models, which may, therefore, provide improved water fluxes and reduce uncertainty of the estimations.

Plain Language Summary The availability of soil water retention data is essential for quantifying mass and energy exchange processes at the interface between land surface and atmosphere. In large-scale applications, soil water retention characteristics usually are estimated with empirical models that, unfortunately, use nonuniform predictors and were developed on subsets of the global distribution of soils. Their reliability for global estimates is often unknown. Using a global database, we developed an ensemble of up to 13 previously published models allowing estimates of soil water retention data under data-poor to data-rich conditions. High-resolution global maps of key points in soil water retention characteristics (and their uncertainties) were produced. These maps suggest that middle and high latitudes in the Northern Hemisphere have larger variability of the estimates. The new model provides more accurate estimates than models currently being used in Earth system models.

1. Introduction

Soils play a fundamental role in mass and energy exchange processes among land surfaces, groundwater, rivers, and the atmosphere (Bittelli et al., 2015; Michael & Cuenca, 1994). Quantification of surface runoff, soil-vegetation-atmosphere transfer fluxes, groundwater recharge, surface energy balances, and land surface temperature must, therefore, rely on correctly parametrized soil hydraulic properties such as soil water retention and hydraulic conductivity characteristics (Chaney et al., 2016; Montzka et al., 2017; Verhoeft & Egea, 2014; Welty & Zeng, 2018; Zhang & Schaap, 2019; Zhao et al., 2018). The experimental determination of soil hydraulic characteristics is time-consuming, labor extensive, and especially impractical for highly heterogeneous soils in large-scale applications (Dai et al., 2013; Gent et al., 2011; Shangguan et al., 2014, 2013). Instead, soil hydraulic properties are often estimated with pedotransfer functions (PTFs), which are empirical data-driven models that commonly utilize available soil attributes as predictors (e.g., soil texture, bulk density, and organic carbon [OC] content). Because of their utility, PTFs have become indispensable components for predicting the dynamics of moisture content in land surface models and global climate models among many other smaller-scale applications (Van Looy et al., 2017).

Over the past decades, considerable national and international efforts have resulted in many PTFs with a variety of statistical approaches (Pachepsky & Schaap, 2004; Zhang et al., 2018). These models were
calibrated using data collected at local or national scales that did not necessarily represent the diversity of the global population of soils, resulting in PTFs that have been demonstrated to produce biased predictions (Dai et al., 2013; Schaap & Leij, 1998; Vereecken et al., 2016). The use of PTFs directly should, therefore, raise one straightforward question: What is the performance of commonly used PTFs in the context of a global-coverage data set?

Van Looy et al. (2017) found that most Earth system models relied on a single PTF to estimate soil hydraulic properties. The selection of a single PTF may result in statistical bias, underestimation of uncertainty, and overconfidence in predictive capabilities (Neuman, 2003). To attempt to mitigate the bias and to expand the support scale of PTFs, uniformly or variably weighted multimodel ensemble estimates can be pursued. Dai et al. (2013) used uniformly weighted ensemble estimates to produce maps of soil hydraulic parameters for China. Utilizing the benefit of validation data, Guber et al. (2006, 2009), however, demonstrated that uniformly weighted ensembles resulted in degraded field-scale estimates.

The main thrust of the present study is to develop a strategy for optimal weighting of PTFs for multimodel estimates in the context of a soil database with global coverage. Previously, such an effort was considered to be challenging (Kishné et al., 2017; Minasny et al., 2013), not only because the calibration data set used to develop PTFs may be a biased selection of the world population of soils but also because soils themselves exhibit an extreme variability in hydraulic (and other) characteristics. For example, the soils in the Sahara formed under completely different climate conditions than those found in boreal zones, leading to vastly different mineral and OC contents and corresponding hydraulic characteristics (Davidson & Janssens, 2006). We will, therefore, also investigate whether one single set of optimal values would suffice to weight PTF ensemble estimates or whether different weightings must be used for soils stratified/categorized by textural classes, OC content, soil order, and soil temperature regime.

In this study, we identified 13 widely cited PTFs used for soil water retention, which were classified into four groups according to input data requirements. Such a hierarchical grouping is critical because not all input data are available in all practical use cases. The PTFs are first evaluated individually against a well-documented soil database (National Cooperative Soil Survey Characterization [NCSS]) covering the continental United States (87.7% of data) and other regions of the globe (12.3% of data). Optimal weights for the PTFs (for all data and stratifications of different soil characteristics, e.g., United States Department of Agriculture [USDA] textural classes, soil OC content, soil orders, and soil temperature) are assigned by minimizing the misfit between multimodel estimates and observed water retention data. The ensemble models are subsequently validated against an independent subset of the database. Finally, global maps of soil water retention data and their uncertainty are produced, which may be useful for a variety of purposes. In addition, we are able to estimate the uncertainty of the multimodel estimates, which will guide further research on improved global-scale PTF estimates.

2. Material and Methods
2.1. PTFs and Soil Hydraulic Functions

Summary characteristics for the 13 PTFs selected for this study appear in supporting information Table S1; R code implementing these models appears in supporting information Code S1. Criteria used to select the PTFs were primarily based on their popularity (as indicated by the number of citations listed in Table S1), but the size of the data set used for calibration of the PTFs was also a secondary consideration, although some PTFs used minimal samples, such as the Campbell and Shiozawa (1992) PTF. PTFs that are soil-specific, or do not estimate a parametric water retention function, were not considered in this study.

There are a number of ways by which PTFs can be grouped and distinguished. First of all, six of the PTFs were derived from two publications (three each from Cosby et al., 1984, and Zhang & Schaap, 2017) and represent different approaches to establish the PTFs (e.g., class PTFs and regression techniques). Second, owing to attempts to construct large representative databases, considerable overlap in calibration data exists among PTFs (see supporting information Text S1 for further details). More recent PTFs often use data that was used for “older” PTFs. Third, all the PTFs predict parameters of different water retention functions. Five PTFs estimate the parameters of Brooks and Corey (1964) water retention model or its Campbell (1974) and Clapp and Hornberger (1978) variants; the remaining eight PTFs estimate
parameters of van Genuchten (1980) model. Although the functional form of the retention equation is relevant, the present work is unable to address this due to the limited number of capillary pressures available in the global-coverage data set used for evaluation (see section 2.2 for more detail).

A key distinction among the PTFs is their predictor variable requirements, which can be sorted as follows. Group A requires only USDA soil textural class (i.e., Carsel & Parrish, 1988; Clapp & Hornberger, 1978; Cosby0, Cosby et al., 1984; and Rosetta3-H1w, Zhang & Schaap, 2017). Group B utilizes soil textural percentage values as predictors (i.e., Cosby1, Cosby et al., 1984; Cosby2, Cosby et al., 1984; and Rosetta3-H2w, Zhang & Schaap, 2017). Group C requires additional soil bulk density (Campbell & Shiozawa, 1992; Rawls & Brakensiek, 1985; and Rosetta3-H3w, Zhang & Schaap, 2017), while Group D further requires soil OC content (Vereecken et al., 1989; Weynants et al., 2009; Wösten et al., 1999). We refer the reader to the references for detailed descriptions of the PTFs.

2.2. Data Set Used for Evaluation and Ensemble Development

The NCSS Database (National Cooperative Soil Survey, 2017) was used to independently evaluate the 13 PTFs and to establish weights for the multimodel ensemble (a brief discussion of other soil databases can be found in supporting information Text S2). After data quality analysis (see Text S2), 49,855 records (having 118,599 water retention points) were selected for use in this study. The World Soil Information Service (WoSIS) data set (Batjes et al., 2017), which includes the NCSS data set and has a much wider spatial soil distribution with over 31 million soil records, was also considered. However, because of inconsistent bulk densities (measured at −0.33 bar and oven dryness) and moisture contents (volumetric and gravimetric) in the WoSIS data, the evaluation is more prone to systematic source-dependent bias, as discussed in supporting information Text S2. Our main focus, therefore, was the more consistent NCSS data set, but a combination of NCSS and WoSIS data was evaluated in section 3.2. The geographical location of the selected NCSS samples and additional WoSIS data is shown in Figure 1a, with 43,650 records from the continental United States and 10,424 from elsewhere.

2.3. Multimodel Ensemble Predictions and Bootstrap Resampling

The weights of multimodel ensemble simulation are determined by minimizing the following:

\[
\chi^2(\mathbf{a}) = \mathbf{z}(\mathbf{z})^T \quad (1)
\]

where \( \mathbf{z} = \frac{1}{N_m} \sum_{j=1}^{N_m} a_j (\mathbf{\Theta}' - \mathbf{\Theta})^2 \quad (2) \)

and \( \mathbf{a} \) is a vector of weights, \( a_j \), for corresponding PTF model \( M_j (j = 1 \ldots N_m) \), \( 0 \leq a_j \leq 1 \); \( N_m \) is the number of PTF ensembles (13 when all PTFs are considered, but 3 or 4 for Groups A through D); and \( \mathbf{\Theta} \) and \( \mathbf{\Theta}' \) are vectors of measured and estimated moisture content, the length of which is the size of calibration samples selected by the bootstrap resampling process (see below). A flowchart of the proposed method can be found in Figure S1. Different algorithms were evaluated for minimization of (1), and we found that a genetic algorithm (Scrucca, 2013; implemented in the statistical package R, Version 3.4.4, see Venables & Smith, 2003) was the most effective method to optimize \( \mathbf{a} \).

The optimization of the ensemble weighting vectors for the entire data set and the four stratifications was coupled with bootstrap (with 100 replicas) resampling (Efron & Tibshirani, 1993) to obtain the uncertainty of \( \mathbf{a} \), which further enabled us to consider evaluation criteria (see below) for independent calibration and validation data (Zhang & Schaap, 2017).

2.4. Evaluation Criteria

The estimated moisture content \( \mathbf{\Theta}'(\psi_i) \) at pressure head \( \psi_i \) is calculated as

\[
\mathbf{\Theta}'(\psi_i) = f(M_j(\mathbf{D}); \psi_i) \quad (3)
\]

where \( \mathbf{D} \) is a vector of predictors specific for PTF model \( M_j \); \( f \) is a water retention function relevant to \( M_j \) (see section 2.1 for list of models), evaluated at pressure head \( \psi_i \).

The criteria used to evaluate different PTFs is to use root-mean-square error (RMSE) of moisture content, defined as
\[ \text{RMSE} = \sqrt{\frac{1}{N_d} \sum_{i=1}^{N_d} (\theta_i - \hat{\theta}_i)^2} \]  

where \( N_d \) is the number of measured \( \theta_i \) and estimated \( \hat{\theta}_i \) moisture contents, for example, 118,599 for the entire data set. Mean error (ME), \( R^2 \), and model selection criteria, such as AIC (Akaike, 1974) and...
AICc (Hurvich & Tsai, 1989), defined in supporting information Text S3, were used to rank and compare individual PTFs and ensemble models. Lower AIC or AICc value indicates a better performing model.

2.5. Optimization of Weights by Stratification of Different Soil Characteristics

The subset of the NCSS data set is not representative of the actual distribution of soils on Earth. For example, there are very limited tropical soils, and data from boreal areas (e.g., Siberia) are missing. There is even less of a guarantee that the original calibration data used to establish the 13 PTFs are representative. By stratifying the selected NCSS data by soil textural class, soil OC content, soil order, and mean soil temperature, it is possible to evaluate whether better estimates can be made by reoptimizing \( a \) in Equation 1 when the data are stratified according to these variables. The maps of soil OC content, soil orders, and soil temperature within the United States are shown in Figures S2a–S2c. NCSS data from other continents were included in the analysis but are not shown for clarity.

2.6. Development of Global Maps of Moisture Content at Specified Pressure Head

Based on the developed ensemble PTFs, global maps of soil water retention were produced by utilizing global soil basic properties as input for PTFs. Here we employed the OpenLandMap data set (see Hengl, 2018a, 2018b, 2018c, 2018d, and Hengl & Wheeler, 2018) as input. OpenLandMap is an updated version of the SoilGrids effort (Hengl et al., 2014, 2017) and provides highly accurate, high-resolution global maps of soil texture, bulk density, and so forth, derived by automated soil mapping and deduced from remote sensing products linking soil properties with various environmental covariates (Luo et al., 2016; Montzka et al., 2017). OpenLandMap data were considered in coarse (10 km) and fine (250 m) resolution. Although it is possible to generate maps of moisture content for arbitrary pressure heads, we chose three key values: moisture content at saturation (i.e., the maximum moisture content at 0 bar), \(-0.33\) bar, and \(-15\) bar. Moisture content at \(-0.33\) bar is commonly used to indicate field capacity, which is the moisture content when gravitationally induced soil-internal drainage of water is minimal; moisture content at \(-15\) bar is associated with wilting point, where most vegetation ceases to extract water from soils (Dane & Topp, 2002; Jury & Horton, 2004; Klute, 1986). The difference between saturation and field capacity has previously been used to estimate saturated hydraulic conductivity (Ahuja et al., 1989; Zhang & Schaap, 2019) and provides a path forward to estimate unsaturated hydraulic conductivity (Schaap & van Genuchten, 2006).

3. Results and Discussion

3.1. Evaluation of Individual PTFs

When considering the performance of the 13 PTFs individually using NCSS data, we found that RMSE is the largest (0.0987 cm\(^3\)/cm\(^3\)) for the Carsel and Parrish PTF and lowest (0.0555 cm\(^3\)/cm\(^3\)) for the Weynants PTF (Table 1). Group-level performance is expected to improve when more predictors are used, though this is not always the case. For example, the Vereecken PTF (Group D) has an RMSE of 0.0658 cm\(^3\)/cm\(^3\), whereas the simple class PTF of Cosby0 (Group A) has a lower RMSE (0.0624 cm\(^3\)/cm\(^3\)). \( R^2 \), AIC, and AICc criteria yield results similar to those shown by RMSE for individual PTFs (shown in supporting information Table S2) and therefore will not be discussed here.

The notable outlier in Table 1 is the Carsel and Parrish PTF, which has the worst performance but is also one of the most widely cited in the literature (Table S1). The poor performance of this model is likely due to its heuristic transformation of Brooks and Corey (1964) retention parameters into van Genuchten (1980) parameters. The Carsel and Parrish PTF is based on the Rawls and Brakensiek (1985) PTF (see Carsel & Parrish, 1988), which has much better performance (0.0629 vs. 0.0987 cm\(^3\)/cm\(^3\)). ME values (also in Table S2) suggest that the Carsel and Parrish PTF has the largest systematic bias (\(-0.0659\)), while Cosby0 has the lowest (\(-0.0009\)). Nearly all PTFs underestimate the observed moisture contents. It is also noted that the interpretation of ME values should be treated with caution because positive and negative values at different pressure heads will tend to cancel out.

3.2. Group Ensembles

Summary results for the optimizations of Groups A through D ensembles as well as those for all 13-PTF models appear in Table 1. Detailed results for each replica and ensemble member are available in supporting information Tables S3–S7. The supplemental results exhibit small differences in RMSE values for the
| Group | RMSE of individual model | Weights of group ensemble model | RMSE of group ensemble model | Weights of overall ensemble model | RMSE of overall ensemble model |
|-------|--------------------------|-------------------------------|-----------------------------|-----------------------------------|-------------------------------|
| Group A PTF (USDA soil texture class) | Cosby0 0.0624 0.0667 | 0.0620 0.0629 | 0.0601 | 0.0536 | 0.0517 |
| Group B PTF (USDA texture percentage: sand, silt, and clay) | Cosby1 0.0681 0.147 | 0.0620 0.0629 | 0.0659 | 0.0528 | 0.0517 |
| Group C PTF (as Group B, but with additional bulk density) | Cosby2 0.0881 0.5961 | 0.0629 0.0646 | 0.0589 | 0.0528 | 0.0517 |
| Group D PTF (as Group C, but with additional soil organic carbon) | Rosetta3 0.0590 0.5392 | 0.0629 0.0646 | 0.0589 | 0.0528 | 0.0517 |

Note: Corresponding mean weights of the ensemble models are also shown. The ensemble models are optimized based on NCSS data set (n = 118,999; optimizations are mean values based on 100 bootstrap replicates). Different groups (A, B, C, and D) are classified based on input data. See text and supporting information Table S1 for detailed explanations of the evaluated PTFs.
calibration and validation members of each bootstrap replica, indicating that each ensemble model is stable. The standard deviation of weights for individual PTFs in each ensemble was less than 0.02; the sum of ensemble weights was always equal to 1.

$\text{RMSE}$ values of ensemble models are all smaller than those of individual ensemble members, decreasing monotonically from 0.0601 (Group A) to 0.0528 (Group D), declining further to 0.0517 cm$^3$/cm$^3$ (full 13-member ensemble). It is worth noting that there is a substantial improvement from Group B to Group C (0.0053 cm$^3$/cm$^3$) but only a small improvement from Group C to Group D (0.0008 cm$^3$/cm$^3$). Bulk density provides information about total porosity but is negatively correlated with soil OC content (Heuscher et al., 2005; Zacharias & Wessolek, 2007; Zhang & Schaap, 2019), which is likely the reason of the small improvements. It is noted that the presented weights are based on all available retention data in selected NCSS records, while optimization using different subsets of retention points yields different weighting for different PTFs (See Text S4 and Figure 1b).

We note that the Rosetta3-H3w, Wösten, and Weynants PTFs carry roughly equal weights (0.17 to 0.20 in the 13-member ensemble), while the Clapp and Hornberger PTF has a surprisingly large contribution (0.14), given its comparatively large individual $\text{RMSE}$ and its simple USDA textural class input requirements. The remaining 30% of the ensemble weights were carried by the nine other PTFs. The results suggest that there is merit in pursuing multimodel ensemble predictions, and the full 13-member ensemble should be considered where possible.

$\text{RMSE}$ values and corresponding weights remain consistent when the analysis is repeated for the extended data set (i.e., NCSS data with additional WoSIS data to increase global coverage; see Table S8). The similarity indicates that although NCSS data are somewhat limited in global geographical scope (most of the data are from the United States), it remains applicable to global scales, while achieving lower $\text{RMSE}$. The extended data set was further separated into the continental United States, tropical regions (from 25°S to 25°N), and the remaining regions based on geographical locations. Corresponding results are also shown in supporting information Table S8 and Figure S3 and are analyzed in Text S5.

### 3.3. Stratification of Different Soil Characteristics

Summary results shown in Figures 1c–1f (with detailed bootstrap-level data appearing in supporting information Tables S9–S12) indicate that PTF weights indeed change for the different stratifications of the data. When stratified by USDA textural class (Figure 1c), the weight for all PTFs varies substantially (except Cosby0 and Carsel and Parrish which remain low), which we believe is caused by data selection effects when the original PTFs were calibrated. This causes some PTFs to outperform others for some parts of the textural triangle but become inferior elsewhere. Different weighting vectors are also obtained when the data are stratified by soil OC content (Figure 1d) or soil order (Figure 1e). When the data are stratified by taxonomic temperature regime (Figure 1f), the PTF weights exhibit less variation.

The results in Figures 1c–1f indicate the likelihood that better estimates can be obtained when the weights are determined for different stratifications of the data; the $\text{RMSE}$ is reduced from 0.0517 to 0.0504, 0.0504, 0.0498, and 0.0483 cm$^3$/cm$^3$ for textural class, OC content, soil order, and taxonomic temperature stratified models, respectively (see Table S2). We note here that the soil order and soil temperature stratified models could only be conducted for a reduced number of samples (91,303 and 80,223 samples, respectively). Strictly speaking, the $\text{RMSE}$ values for soil order and temperature regime cannot be compared accurately to that of the 13-member ensemble. However, $\text{AIC}$ and $\text{AICc}$ values confirm that the improvements by reoptimizing the PTF weights for different stratified models are statistically significant (see Table S2).

The performance of ensemble models varies among each category in different stratifications. We show the ensemble model performance relative to the best individual PTF and the performance of the 13 PTFs individually for soils in different regions, pressure head, textural class, OC content, soil order, and taxonomic temperature, respectively. We recommend that the appropriate ensemble model(s) should be used where possible, but it is possible that individual PTFs may outperform these in selected cases (see Text S7 and Table S13).

10.1029/2020GL088819
3.4. Global Maps of Moisture Content at Saturation and at $-0.33$ and $-15$ bar

To demonstrate the credibility of the predicted global maps, we carried out a test with NCSS data, which provide moisture content at $-0.33$ and $-15$ bars at georeferenced points (observed saturated moisture content is unfortunately not available). To this end, we retrieved the coarse and fine OpenLandMap estimates at these points and ran the multimodel ensembles. Comparison of OpenLandMap estimates with the observed NCSS data showed that the high-resolution OpenLandMap (250 m) is substantially better than the low-resolution estimate (see Figure S4). In the case of moisture content at $-0.33$ bar, $R^2$ values increase from 0.2099 to 0.4957 when 250 m OpenLandMap data are used instead of 10 km data. For the moisture content at $-15$ bar, $R^2$ increases from 0.3013 to 0.6465. These results are promising considering a tenfold cross validation of observed and estimated 250 m SoilGrids data (Hengl et al., 2017) produced $R^2$ between 0.63 and 0.79, depending on the soil variables considered (sand, silt, clay, bulk density, and OC); a corresponding analysis for OpenLandMap data is not available. These results further suggest that estimated global moisture maps depend not only on the PTF models but also substantially on the accuracy and resolution of the maps of input data. Artifacts present in the input data of PTFs (i.e., different resolution data set of OpenLandMap or SoilGrids data) will lead to artifacts in the predicted maps.

We also show the comparison of moisture content at $-0.33$ and $-15$ bars estimated from overall ensemble PTFs and OpenLandMap data set (Hengl & Gupta, 2019) in Figures S5 and S6. OpenLandMap data set provides more accurate estimation compared with ensemble PTFs, since the former are estimated directly from observations based on machine learning. In contrast, the latter are based on PTFs indirectly from soil properties, such as soil texture and bulk density. However, ensemble PTFs can not only provide estimations of moisture content at $-0.33$ and $-15$ bar but moisture content at any pressure head and the entire soil water retention curve. Please see the discussion of the comparison of the two data sets in Text S6.

Figure 2 shows the mean values and the corresponding coefficient of variation (CV) of the 13-member ensemble estimates. Maps of ensemble estimates and CV values of saturated moisture content and moisture content at $-0.33$ and $-15$ bar for Groups A through D and the 13-member ensemble estimates can be found.
in Figures S7–S9. The CV values represent the ratio of the standard deviation to the mean of the 100-member bootstrap estimates for each grid location. We utilized the 250 m resolution OpenLandMap data but plotted the maps in the center of the 10 km pixel for display reasons since 250 m data cannot be represented in a meaningful way in a paper or digital copy of the present work.

Figure 2 shows that low saturated moisture content values occur in most of the Southern Hemisphere and low latitudes and midlatitudes of the Northern Hemisphere, while high values are found in high latitudes of the Northern Hemisphere. Low values of moisture content at −0.33 and −15 bar are found in the Sahara and Arabian Peninsula, while high moisture content at −0.33 bar is shown in Canada, most of Siberia, and part of South and Southeast Asia. High values of moisture content at −15 bar are found in Mexico, Central America, South and Southeast Asia, and parts of South America and parts of tropical Africa. Compared to Groups A and B, Groups C and D ensembles exhibit substantially higher saturated moisture contents and moisture content at −0.33 bar in parts of Canada and Siberia (see Figures S7 and S8); these higher estimates persist in the 13-member ensemble (Figures S7i and S8i). This is presumably due to the information provided by bulk density and OC content, used by Groups C and D ensembles. Soils in these regions are known to have high OC contents, which are inversely correlated with bulk density (Zacharias & Wessolek, 2007).

There are also distinct differences between the maps for moisture content at −15 bar produced by Groups A, B, and C ensembles and those generated with the Group D and the full 13-member ensemble (Figure S9). See our discussion of the distinct differences and the variation of CV values in supporting information Text S8. High values of CV might serve as a guide where PTF improvement should be conducted.

4. Summary and Conclusions

Our work leads to five major points:

1. Thirteen widely used PTFs for the estimation of water retention characteristics were evaluated on independently measured moisture contents available in a large soil data set with global coverage (NCSS and WoSIS data set).
2. Thirteen PTFs were grouped into four classes according to input data requirements. Weighted multimodel ensemble estimates for each PTF group resulted in improved performance relative to group member PTFs. A further improvement was achieved by a weighted ensemble of all 13 models.
3. Model weights changed for stratifications of different soil characteristics, that is, USDA textural class, OC content, soil order, and soil temperature classification, which indicates that each of the PTFs had intrinsic biases. Improved ensemble PTFs can be obtained when the weights are determined for different stratifications of the data, and further research remains necessary to derive a metadata PTF that effectively deals with soil-dependent weights.
4. Validation of moisture content estimated from the ensemble models applied to the global maps against the observed NCSS data set suggested that produced moisture maps depend not only on the PTF models but also heavily on the accuracy of the soil maps but shows promising results on high resolution of basic soil property data set. Maps of moisture content at saturation and at −0.33 and −15 bar pressure head (and their uncertainties) were derived for the ensemble models. The maps produced in the present study have a resolution of 10 km; higher-resolution maps and data structures with complete retention curves (and associated uncertainty) can also be generated.
5. The ensemble PTFs provide estimates that are superior over PTFs currently being used in Earth system modeling. Use of estimates by weighted PTF ensembles may, therefore, provide more accurate estimates of moisture content and water flux and reduce statistical bias and reduce uncertainty of estimations for soil water balance models, hydrological and ecological models, crop growth models, land surface models, weather forecast models, air quality models, and global climate models.

Data Availability Statement

The code and global maps (in GeoTIFF format) of mean values and coefficients of variation of saturated moisture content, field capacity, and permanent wilting point estimated from overall 13-PTF ensemble model can be downloaded online (from https://doi.org/10.7910/DVN/VPIN2B).
Acknowledgments
Zhang is supported by the National Natural Science Foundation of China (Grant 41807181). Schaap was supported by the USDA National Institute of Food and Agriculture, Hatch Project W-3188: Soil, Water, and Environmental Physics Across Scales (https://www.nimss.org/projects/view/mrp/outline/16636). We thank the seven reviewers and two editors, Valeriy Ivanov and Bayani Cardenas, for their insightful comments and suggestions.

References
Ahuja, L. R., Cassel, D. K., Bruce, R. R., & Barnes, B. B. (1989). Evaluation of spatial distribution of hydraulic conductivity using effective porosity data. Soil Science, 148(6), 404–411. https://doi.org/10.1007/BF00010649-1989-120000-00002
Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716–723. https://doi.org/10.1109/tac.1974.110076
Batjes, N. H., Ribeiro, E., van Oostrom, A., Leenaars, J., Hengl, T., & Mendes de Jesus, J. (2017). WoSIS: Providing standardized soil profile data for the world. Earth System Science Data, 9(1), 1–14. https://doi.org/10.5194/essd-9-1-2017
Bittelli, M., Campbell, G. S., & Tomet, F. (2015). Soil physics with Python: Transport in the soil-plant-atmosphere system. New York, USA: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199683093.001.0001
Brooks, R., & Corey, A. (1964). Hydraulic properties of porous media. Hydrology Papers, Colorado State University, 3, 1–27.
Campbell, C. G., & Shiozawa, S. (1992). Prediction of hydraulic properties of soils using particle-size distribution and bulk density data. In Indirect methods for estimating the hydraulic properties of unsaturated soils (pp. 317–326). Riverside: University of California.
Campbell, G. S. (1974). A simple method for determining unsaturated conductivity from moisture retention data. Soil Science, 117(6), 311–314. https://doi.org/10.1007/BF00741901
Casel, 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
Chaney, N. W., Herman, J. D., Ek, M. B., & Wood, E. F. (2016). Deriving global parameter estimates for the Noah land surface model using FLXNET and machine learning. Journal of Geophysical Research, 121, 13,218–13,235. https://doi.org/10.1002/2016JD024821
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
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(6), 682–690. https://doi.org/10.1029/WR020i006p00682
Dai, Y., Shangguan, W., Duan, Q., Liu, B., Fu, S., & Niu, G. (2013). Development of a China dataset of hydraulic parameters using pedotransfer functions for land surface modeling. Journal of Hydrometeorology, 14(3), 869–887. https://doi.org/10.1175/JHM-D-12-0134.1
Dane, J. H., & Topp, G. C. (Eds) (2002). Methods of soil analysis, part 1, physical and mineralogical methods (second edition). Madison, Wisconsin, USA: Soil Science Society of America.
Guber, A. K., Pachepsky, Y. A., van Genuchten, M. T., Rawls, W. J., Simunek, J., Jacques, D., et al. (2006). Field-scale water flow simulations using ensembles of pedotransfer functions for soil water retention. Vadose Zone Journal, 5(1), 234–247. https://doi.org/10.2136/vzj2005.0111
Guber, A. K., Pachepsky, Y. A., van Genuchten, M. T., Simunek, J., Jacques, D., et al. (2009). Multimodel simulation of water flow in a field soil using pedotransfer functions. Vadose Zone Journal, 8(1), 1–10. https://doi.org/10.2136/vzj2007.0144
Hengl, T. (2018a). Clay content in % (kg/kg) at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution [data set]. 10.5281/ZENODO.2525663
Hengl, T. (2018b). Sand content in % (kg/kg) at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution [data set]. 10.5281/ZENODO.2525662
Hengl, T. (2018c). Silt content in % (kg/kg) at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution [data set]. 10.5281/ZENODO.2525676
Hengl, T. (2018d). Soil bulk density (fine Earth) 10 x kg/m-cubic at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution [data set]. 10.5281/ZENODO.2525665
Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B. M., Ribeiro, E., et al. (2014). SoilGrids1km—Global soil information based on automated mapping. PLoS ONE, 9(8), e105992-e105917. https://doi.org/10.1371/journal.pone.0105992
Hengl, T., & Gupta, S. (2019). Soil water content (volumetric %) for 33 kPa and 1500 kPa suction predictions at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution [data set]. 10.5281/ZENODO.2784001
Hengl, T., Mendes De Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzales, M., Kilibarda, M., Blagotic, A., et al. (2017). SoilGridds250m: Global gridded soil information based on machine learning. PLoS ONE, 12(2), e0169748. https://doi.org/10.1371/journal.pone.0169748
Hengl, T., & Wheeler, I. (2018). Soil organic carbon content in x g/kg at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution (Version v0.2) [data set]. 10.5281/ZENODO.2525553
Heuscher, S. A., Brandt, C. C., & Jardine, P. M. (2005). Using soil physical and chemical properties to estimate bulk density. Soil Science Society of America Journal, 69(6), 1943–1954. https://doi.org/10.2136/sssaj2005.0051
Hurvich, C. M., & Tsai, C. L. (1989). Regression and time series model selection in small samples. Biometrika, 76(2), 297–307. https://doi.org/10.1093/biomet/76.2.297
Jury, W. A., & Horton, R. (2004). Soil physics (6th ed.). New York: John Wiley & Sons, Inc.
Kishnė, A. S., Yamam, Y. T., Morgan, C. L. S., & 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
Khun, A. (Ed) (1986). Methods of soil analysis, part 1, physical and mineralogical methods (second edition) (Vol. 5, pp. 87–89).
Madison, Wisconsin, USA: American Society of Agronomy, Inc: Soil Science Society of America, Inc. https://doi.org/10.1002/geo.340050110
Luo, Y., Ahlström, A., Allison, S. D., Batjes, N. H., Brovkin, V., Carvalhais, N., et al. (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
Michael, E. K., & Cuenca, R. H. (1994). Variation in soil parameters: Implications for modeling surface fluxes and atmospheric boundary-layer development. Boundary-Layer Meteorology, 70(4), 369–383. https://doi.org/10.1007/bf00731776
Minaewy, B., Whelan, B. M., Tsiantakis, L., & McIntyre, A. B. (2013). Pedometrics research in the vadose zone—Review and perspectives. Vadose Zone Journal, 12(4), 1, v2012.0141-20. https://doi.org/10.2136/vzj2012.0141
Montzka, C., Herbst, M., Weihermüller, L., Verhoef, A., & Vereecken, H. (2017). 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(2), 529–543. https://doi.org/10.5194/essd-2017-13

National Cooperative Soil Survey. (2017). National Cooperative Soil Survey Characterization database. Retrieved January 1, 2017, from http://ncslabdatamart.sc.egov.usda.gov

Neuman, S. P. (2003). Maximum likelihood Bayesian averaging of uncertain model predictions. Stochastic Environmental Research and Risk Assessment, 17(5), 291–305. https://doi.org/10.1002/ser.770

Pachepsky, Y. A., & Schaap, M. G. (2004). Development of pedotransfer functions in soil hydrology. In Y. A. Pachepsky & M. G. Schaap (Eds.), Developments in soil science (Vol. 30, pp. 1–512). Amsterdam, The Netherlands: Elsevier. https://doi.org/10.1016/b978-0444-24810-4/00102-7

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

Schaap, M. G., & Leij, F. J. (1998). Database related accuracy and uncertainty of pedotransfer functions. Soil Science, 168(10), 765–779. https://doi.org/10.1007/s0010694-199810000-00001

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), 27–34. https://doi.org/10.2136/vzj2005.0005

Scurra, L. (2013). GA: A package for genetic algorithms in R. Journal of Statistical Software, 44(2), 1–26. https://doi.org/10.18637/jss.v044.i02

Shangguan, W., Dai, Y., Duan, Q., Liu, B., & Yuan, H. (2014). A global soil data set for Earth system modeling. Journal of Advances in Modeling Earth Systems, 6(1), 249–263. https://doi.org/10.1002/2013MS000293

Shangguan, W., Dai, Y., Liu, B., Zhu, A., Duan, Q., Wu, L., et al. (2013). A China data set of soil properties for land surface modeling. Journal of Advances in Modeling Earth Systems, 5(2), 212–224. https://doi.org/10.1002/jame.20026

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), 892–899. https://doi.org/10.2136/sssaj1980.03615995004400050002x

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

Wageningen. 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), 892–899. https://doi.org/10.2136/sssaj1980.03615995004400050002x

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/10.1007/bf00106948-198912000-00001

Vereecken, H., Schnepf, A., Hopmans, J. W. W., Javou, M., Or, D., Reeve, T., et al. (2016). Modeling soil processes: Review, key challenges, and new perspectives. Vadose Zone Journal, 15(5), 795, v015.09.0131–820. https://doi.org/10.2136/vzj2015.09.0131

Verhoef, A., & Ega, G. (2014). Modeling plant transpiration under limited soil water: Comparison of different plant and soil hydraulic parameterizations and preliminary implications for their use in land surface models. Agricultural and Forest Meteorology, 191, 22–32. https://doi.org/10.1016/j.agrformet.2014.02.009

Welty, J., & Zeng, X. (2018). Does soil moisture affect warm season precipitation over the southern Great Plains? Geophysical Research Letters, 45, 7866–7873. https://doi.org/10.1002/2018GL078598

Weynants, M., Vereecken, H., & Javou, M. (2009). Revisiting Vereecken pedotransfer functions: Introducing a closed-form hydraulic model. Vadose Zone Journal, 8(1), 86–95. https://doi.org/10.2136/vzj2008.0062

Winter, J. 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

Zacharias, S., & Wessoek, 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

Zhang, Y., & Schaap, M. G. (2016). Weighted recalibration of the Rosetta pedotransfer model with improved estimates of hydraulic parameter distributions and summary statistics (Rosseta3). Journal of Hydrology, 347, 39–53. https://doi.org/10.1016/j.jhydrol.2017.01.004

Zhang, Y., & Schaap, M. G. (2019). Estimation of saturated hydraulic conductivity with pedotransfer functions: A review. Journal of Hydrology, 575, 1011–1030. https://doi.org/10.1016/j.jhydrol.2019.05.058

Zhang, Y., Schaap, M. G., & Zha, Y. (2018). A high-resolution global map of soil hydraulic properties produced by a hierarchical parameterization of a physically-based water retention model. Water Resources Research, 54, 9774–9790. https://doi.org/10.1029/2018wr023539

Zhao, H., Zeng, Y., Iv, S., & Su, Z. (2018). Analysis of soil hydraulic and thermal properties for land surface modeling over the Tibetan Plateau. Earth System Science Data, 10(2), 1031–1061. https://doi.org/10.5194/essd-10-1031-2018

References From the Supporting Information

Batjes, N. H. (2000). Global data set of derived soil properties, 0.5-degree grid (ISRIC-WISE). Oak Ridge, TN: ORNL DAAC. https://doi.org/10.3334/ORNLDAAC/546

Batjes, N. H. (2009). Harmonized soil profile data for applications at global and continental scales: Updates to the WISE database. Soil Use and Management, 25(2), 124–127. https://doi.org/10.1111/j.1747-7733.2009.00202.x

Batjes, N. H., Ribeiro, E., & Van Oostrom, A. (2020). Standardised soil profile data to support global mapping and modelling. Earth System Science Data, 12(1), 299–320. https://doi.org/10.5194/essd-12-299-2020

Carrera, J., & Neuman, S. P. (1986). Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood method incorporating prior information. Water Resources Research, 22(2), 199–210. https://doi.org/10.1029/WR022i002p00199

de Lannoy, G. 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/10.1029/2014ms00330

Leenaars, J. G. B. (2013). Africa Soil Profiles Database—A Compilation of georeferenced and standardised legacy soil profile data for Sub Saharan Africa (Version 1.1), Africa Soil Information Service (AFSIS) and ISRIC—World Soil Information. Arroyo et al. (eds.), 2014b. Wageningen.

Seber, G. A. F., & Lee, A. J. (2003). Linear regression analysis (second edition). Hoboken, NJ, USA: John Wiley & Sons, Inc.. Retrieved from https://doi.org/10.1002/9780471722199.ch3
Seber, G. A. F., & Wild, C. J. (1989). *Nonlinear regression*. StatBook. Hoboken, NJ: John Wiley & Sons, Inc. https://doi.org/10.1002/0471725315

Van Engelen, V. W. P. (2011). Standardizing soil data (e-SOTER regional pilot platform as EU contribution to a global soil information system). In *International innovation*. June, 48-49

Ye, M., Meyer, P. D., & Neuman, S. P. (2008). On model selection criteria in multimodel analysis. *Water Resources Research*, 44, W03428. https://doi.org/10.1029/2008WR004803