Choice of Pedotransfer Functions matters when simulating soil water balance fluxes

L. Weihermüller*[1,7], P. Lehmann[2], M. Herbst[1], M. Rahmati[3], A. Verhoef[4], D. Or[2,5], D. Jacques[6], and H. Vereecken[1,7]

[1] Agrosphere Institute IBG 3, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany,
[2] Department of Environmental Systems Science, ETH Zurich, Switzerland
[3] University of Maragheh, Department of Soil Science and Engineering, Iran
[4] University of Reading, Department of Geography and Environmental Science, Reading, UK
[5] Division of Hydrologic Sciences (DHS) - Desert Research Institute, Reno, NV, USA
[6] Belgian Nuclear Research Centre SCK CEN, Institute for Environment, Health and Safety (EHS), Mol, Belgium
[7] International Soil Modeling Consortium (ISMC), Forschungszentrum Jülich GmbH, Jülich, Germany

*Corresponding author: Lutz Weihermüller, Forschungszentrum Jülich IBG-3 Agrosphere Institute D-52425 Jülich, Voice: ++49 (0)2461 61 8669 email: l.weihermueller@fz-juelich.de

Abstract

Modelling of the land surface water-, energy-, and carbon balance provides insight into the behaviour of the Earth System, under current and future conditions. Currently, there exists a substantial variability between model outputs, for a range of model types, whereby differences between model input parameters could be an important reason. For large-scale land surface, hydrological, and crop models, soil hydraulic properties (SHP) are required as inputs, which are estimated from pedotransfer functions (PTFs). To analyse the functional sensitivity of widely used PTFs, the water fluxes for different scenarios using HYDRUS-1D was simulated and predictions compared. The results showed that using different PTFs causes substantial variability in predicted fluxes. In addition, an in-depth analysis of the soil SHPs and derived soil characteristics was performed to analyse *why* the SHPs estimated from the different PTFs cause the model to behave differently.

The results obtained provide guidelines for the selection of PTFs in large scale models. The model performance in terms of numerical stability, time-integrated behaviour of cumulative fluxes, as well as instantaneous fluxes was evaluated, in order to compare the suitability of the PTFs. Based on this, the Rosetta, Wösten (), and Tóth PTF seem to be the most robust PTFs for the Mualem van Genuchten SHPs and the PTF of Cosby for the Brooks Corey functions. Based on our findings, we strongly recommend to harmonize the PTFs used in model inter-comparison studies to avoid artefacts originating from the choice of PTF rather from different model structures.

Plain Language Summary

Hydrological models need information about the soil physical characteristics (soil hydraulic parameters), which are in general not available if the models are applied at larger scales (region to global scale). Therefore, pedotransfer functions (PTFs) are classically used, which relate easily available soil properties such as sand-, silt-, clay-content, soil organic carbon content, and soil bulk density, which are available from soil maps, to the soil hydraulic parameters. Unfortunately, there are many different PTFs available in literature. In the study presented, we analysed the impact of different PTFs on the simulation results of water fluxes and found, that the choice of PTF impacts the simulation results. Further, some PTFs were identified as being less robust compared to others. In general, the study shows that harmonizing PTFs in model-inter-comparisons is needed to avoid artefacts originating from the choice of PTF rather from different model structures.

Keywords: pedotransfer functions, land surface models, LSM, hydrological models, crop models, model inter-comparison, model ensemble mean

1. Introduction

Water fluxes and soil water content are key variables in the terrestrial system as they control the exchange of water and energy between the land-surface and the atmosphere (e.g., Vereecken et al., 2015). Modelling of the water flow in the unsaturated zone, and the uncertainty in the parameters used to simulate water flow, has been a topic of intense research for many years, both in the soil hydrological and land surface modelling community (Shao & Irannejad, 1999; (Tietje & Tapkenhinrichs, 1993; Vereecken et al., 2008; Iwema et al. (2017)). Moreover, climate modellers have studied the role of soil water content, and strongly related processes such as evapotranspiration, in climate and atmospheric processes (Koster & Suarez 2001; Ek & Holtslag, 2004; van den Hurk et al., 2008; Seneviratne et al., 2010; Groh et al., 2020). In this context, we require reliable estimates of soil hydraulic properties at point to global scale (Cornelis et al., 2001, van Looy et al., 2017). Measuring these properties is tedious, time and cost expensive, and prone to measurement errors. Often, taking measurements is not feasible due to the complexity and/or size of the terrestrial system under investigation. To overcome this problem, pedotransfer functions (PTFs), which estimate these essential soil properties from easily available soil parameters, such as soil texture, soil structure, bulk density, and soil organic matter, have been developed. An extensive overview of existing PTFs was provided by Vereecken et al. (2010) and by van Looy et al. (2017).

Soil properties used as basic input data to estimate the soil hydraulic properties with PTFs are grouped into four categories: (1) soil particle size or soil texture, (2) easily measurable hydraulic properties, (3) morphological properties, and (4) chemical properties (Espino et al., 1995, Vereecken et al. 2010; van Looy et al. 2017; Rahmati et al., 2013; Neyshabouri et al., 2015; Rahmati and Neyshabouri, 2016). In general, two different types of PTF can be distinguished, namely point and parametric PTFs. Point PTFs estimate soil water content (or hydraulic conductivity) values at predefined pressure head values (e.g., field capacity or wilting point), whereas parametric PTFs provide the parameters than can be used in hydraulic functions (water retention curve (WRC) and hydraulic conductivity curve (HCC)) of the Brooks & Corey (1964) or van Genuchten (1980) formulation. The most useful PTFs developed in recent years are the parametric PTFs because they can be used to calculate the WRC and HCC, which are used to simulate the water fluxes in the numerical models. Secondly, PTFs are classified into class and continuous PTFs. Class PTFs are look up tables, where the hydraulic parameters are listed for typical soil textural classes (e.g., 12 USDA soil classes). Continuous PTFs, on the other hand, use mathematical descriptions, e.g., regression

functions, to calculate the hydraulic parameters from the entire range of data inputs like e.g. soil texture.

It has been shown that PTFs are highly accurate for the area (or the input data range) they were developed for, but have limited accuracy if applied outside these regions (Vereecken et al., 2010). Several reviews about the accuracy and reliability of PTFs for the van Genuchten model (VGM) have already been published (e.g., Wösten et al., 2001; Schaap, 2004; Donatelli et al., 2004). Hereby, the predicted hydraulic function from the PTFs were compared to the measured data and the goodness of fit of the prediction was evaluated. The authors used two metrics to determine the performance of the PTF: 1) the term accuracy was related to the comparison between predicted and measured values of water content or hydraulic conductivity that were used to develop the PTF; 2) reliability was related to the evaluation of PTFs on measured values that were different from those that were used to develop the PTFs (Wösten et al., 2001, Zhang et al., 2020). Reliability studies are typically validation studies such as those performed by Tietje & Tapkenhinrichs (1993), Wösten et al. (2001), and Wagner et al. (2001). Despite much progress in developing PTFs and in identifying appropriate PTF predictor candidates, some unresolved or unexplained variability still exists at the level of the soil sample (Schaap & Leij, 1998), which plays and important role when functional aspects of soils are being studied and analysed using numerical models (e.g., Christiaens & Feyen, 2001). Functional aspects already studied are the impact of PTFs on water supply capacity (Vereecken et al., 1992), ground water recharge (Vereecken et al., 1992), and aeration (Wösten et al., 2001). In the study of Vereecken et al. (1992), the authors showed that 90% of the variation in the predicted soil water supply was attributed to estimation errors in hydraulic properties using the PTFs developed by Vereecken et al. (1989, 1990). Chirico et al. (2010), on the other hand, evaluated the effect of PTF prediction uncertainty on the components of the soil water balance at the hillslope scale. One major result was that the simulated evaporation was much more affected by the PTF model error than by errors resulting from uncertainty in the input data (e.g., soil texture).

Land surface models (LSMs), when embedded in numerical weather prediction or climate models, generally operate at large scales (regional, continental to global scales) and rely on PTFs to predict the hydraulic functions needed to solve the Richards equation for the water flow. Different LSMs use different PTFs for this purpose (Vereecken et al., 2019). As Vereecken et al. (2019) showed, not only different PTFs but also different hydraulic models (Campbell, Brooks and Corey, or Mualem-van Genuchten) are in use. Knowing that different PTFs and/or the choice of the hydraulic model will impact the outcome of the water flow

simulations (e.g., Gruber et al., 2006; Yakirevich et al., 2013), a key question is how recently launched LSM inter-comparison activities of the Land Surface Schemes (LSSs) embedded in LSMs, such as those under the Global Energy and Water Exchanges (GEWEX) GLASS project (https://www.gewex.org/panels/global-landatmosphere-system-study-panel/) or model inter-comparisons initiated by the World Climate Research Programme (https://www.wcrp-climate.org/wgcm-cmip/wgcm-cmip6, of which GEWEX is part), such as CMIP6 and its predecessors, will be impacted by the choice of PTF.

Therefore, the aim of this study is to systematically analyse the functional sensitivity to the choice of different PTFs using a physically-based numerical model. As the 'truth' of this model exercise is unknown, the performance of the model runs with its hydraulic parameters derived from a set of individual PTFs will be evaluated against the ensemble mean as best predictor, as well as against the 70 and 90 tolerance intervals of the ensemble range. The numerical exercise is structured in the following way: 1) model runs for homogeneous soil profiles without vegetation, 2) homogeneous soil profile covered with grass and wheat, 3) layered bare soil, 4) layered vegetated soil (grass and wheat), and 5) influence of a fluctuating water table in a layered grass vegetated soil. Finally, additional soil physical properties were calculated based on the estimated soil hydraulic parameters obtain from the PTFs, which were used to explain the differences observed in simulated water fluxes. As some LSMs also use class PTFs (van Looy et al., 2017, Vereecken et al., 2019), we will also analyse the use of this type of PTF, and the associated errors when simulating water fluxes. We formulate three hypotheses 1) the use of different PTFs will lead to systematically different hydrological states and fluxes (e.g., net infiltration, evapotranspiration, root zone water availability, drainage), 2) some PTFs can be identified which perform distinctively differently from the ensemble spread in terms of 90 % tolerance interval outliers, and 3) the differences in predicted states and fluxes simulated with inputs from different PTFs will be reduced with increasing model setup complexity.

2. Materials and Methods

2.1. Hydraulic functions

Three pairs of hydraulic functions are widely used in hydrological modelling, namely those developed by Brooks and Corey (1964), Campbell (1974), and van Genuchten (1980.

The Brooks and Corey (BC) (1964) water retention function is given by:

$$S_e = \begin{cases} |\alpha h|^{-n} & h < -1/\alpha \\ 1 & h \geq -1/\alpha \end{cases} \quad [1]$$

where α is the reciprocal of the air entry value (or bubbling pressure) [cm^{-1}], n is a dimensionless shape parameter [-] (related to $1/b$ for the original Brooks-Corey b parameter), h is the pressure head [cm], and S_e is the effective saturation [-] given by:

$$S_e = \frac{\theta - \theta_r}{\theta_s - \theta_r} \quad [2]$$

where θ is the actual water content [$cm^3\ cm^{-3}$], θ_r is the residual water content [$cm^3\ cm^{-3}$], and θ_s is the saturated water content [$cm^3\ cm^{-3}$].

The unsaturated hydraulic conductivity is given by:

$$K = K_s S_{e^{2/n+3}} \quad [3]$$

where K_s is the saturated hydraulic conductivity [$cm\ d^{-1}$].

The Campbell (1974) water retention function is a modification of that introduced by Brooks and Corey (1964), with θ_r set to 0.

The Mualem van Genuchten function (MvG) (van Genuchten, 1980) is given by:

$$\theta_{(h)} = \theta_r + \frac{\theta_s - \theta_r}{\left(1 + |\alpha h|^n\right)^m} \quad [4]$$

where m is a shape factor related to n by $m = 1\text{-}1/n$.

For the VGM model, the unsaturated hydraulic conductivity is calculated by:

$$K_{(h)} = K_s S_{e^{\lambda}} \left[1 - \left(1 - S_{e^{1/m}}\right)^m\right]^2 \quad [5]$$

where λ is the tortuosity factor [-].

2.2. Pedotransfer Functions

For this study 13 pedotransfer functions (PTFs) were used, whereby eight predict the hydraulic parameters for the MvG (van Genuchten, 1980) and five PTFs predict the parameters in the BC (Brooks & Corey, 1964) or Campbell (Campbell, 1974) functions. Out of these 13 PTFs, four were so called *class-transfer* functions, where only the USDA textural classes will be used as input for the prediction of the hydraulic parameters. It has to be noted that the PTF of Clapp and Hornberger, (1978) (from here on Clapp&Hornberger) does not specify hydraulic parameters for the silt class. All other PTFs use textural information (gravimetric percentage of sand, silt, and clay) as basic inputs. Additionally, some PTFs require information about bulk density (BD) such as the PTFs of Schaap et al. (2001) (here referred to as Rosetta SSC+BD), Wösten et al. (1999) (here Woesten), Weynants et al. (2009) and Weihermüller et al. (2017) (here Weynants), and that of Tóth et al., (2015) for the topsoil (here 'Toth continuous'). Others need information about the organic carbon content (C_{org}),

which are the Woesten, Weynants, and 'Toth continuous' PTFs. Here, it has to be noted that an updated version of Rosetta (Rosetta3) is also available (Zhang & Schaap, 2017), which provides more accurate soil hydraulic parameters compared with the estimation from the original Rosetta model. Nevertheless, we decided to use the older Rosetta version. as it is widely in use and also imbedded in some hydrological software such as HYDRUS (Šimůnek et al., 2008; Šimůnek & van Genuchten, 2008).

Soil organic carbon is used as a predictor for the PTFs as it affects soil bulk density, hydraulic conductivity, and water retention because of its effect on soil structure and adsorption properties (van Genuchten & Pachepsky, 2011).

Total porosity ϕ, as estimated from bulk density (BD), is used only by Rawls and Brakensiek (1985) (here Rawls MvG), and pH and cation exchange capacity (CEC) are inputs in 'Toth continuous'. An overview of all used PTFs, their abbreviations, and their inputs is provided in Tab. 1. The region from where data were taken to train the PTF are from either the USA or Europe. Rosetta is the only PTF combining two data regions, whereas Weynants PTF is based on samples from Belgium only. In addition, the number of samples used for PTF development greatly differs, ranging from 5320 for Rawls PTFs to 166 for Weynants. Important for the PTF development is the data used to generate the PTFs, whereby either only retention data ($\theta_{(h)}$) or a combination of retention ($\theta_{(h)}$) and unsaturated hydraulic conductivity ($K_{(h)}$) data was used. $\theta_{(h)}$ and $K_{(h)}$ were used in the development of the PTFs Rosetta, Woesten, Weynants, and Toth, whereby the percentage of available $K_{(h)}$ data is typically low compared to the availability of $\theta_{(h)}$ data, generally due to the more complex and laborious procedures required to determine $K_{(h)}$. Even though in some cases both types of data ($\theta_{(h)}$ and $K_{(h)}$) were used in the development of some PTFs, the data were not jointly inverted to estimate the hydraulic parameters, meaning that Rosetta, Woesten, and Toth fitted the hydraulic parameters sorely on the retention curve and used the fitted α and n values of the Mulaem van Genuchten equation to predict $K_{(h)}$. In contrast, Weynants used joint inversion of both hydraulic characteristics ($\theta_{(h)}$ and $K_{(h)}$) simultaneously to estimate the parameters including a near saturation hydraulic conductivity K_s^* at a predefined pressure head of – 6 cm. All other PTFs either used the closed form expression of van Genuchten (1980) or Brooks and Corey (1964) to predict $K_{(h)}$, using the estimated parameters from the retention data, together with measured K_s values, to estimate the unsaturated hydraulic conductivity based on either van Genuchten (1980) or Brooks and Corey (1964).

In this study, we will compare model simulations for 12 soil textural classes. For the estimation of hydraulic parameters from texture based continuous PTFs a representative soil

texture was used for each soil class located in the centre of the respective class area in the textural triangle; bulk density and C_{org} were set to 1.4 g cm^{-3} and 1 %, respectively. The texture of the corresponding class is depicted in Fig. 1 and the predicted hydraulic parameters for all applied PTFs are listed in Annex Tab. 1 and Annex Tab. 2.

In general, it is known that relatively small changes in the shape of the soil water retention curve near saturation can significantly affect the results of numerical simulations of water flow for variably saturated soils, including the performance of the numerical stability and rate of convergence (Vogel et al., 2001; Schaap & van Genuchten, 2006). To address this problem, especially in fine textured soils, the estimated air entry value (i.e., the reciprocal of α) from the PTF for the van Genuchten formulation (Eq. 4) was set to -2 cm as proposed by Vogel et al. (2001), whenever the originally proposed set of hydraulic properties from the PTF did not lead to numerical convergence.

2.3. Numerical Modelling

For the simulation of vertical water flow, the one-dimensional Richards equation (Eq. 6) was solved using the finite element code HYDRUS-1D (Šimůnek et al., 2008; Šimůnek & van Genuchten, 2008):

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[K_{(\theta)}\left(\frac{\partial h}{\partial z} + 1\right)\right] - Q \qquad [6]$$

where z represents the vertical coordinate [cm], positive in the downward direction, Q is the source/sink term, θ is the volumetric water content [cm^3 cm^{-3}], and $K_{(\theta)}$ is the unsaturated hydraulic conductivity [cm d^{-1}] as a function of water content.

A 200 cm soil profile was simulated, and the lower boundary condition of the flow domain was defined as free drainage, which is typically used when the ground water table is far below the soil surface. As a second option, a fluctuating ground water table was prescribed as a Dirichlet boundary condition. For the variable ground water table depth, a simple sine curve of the ground water table fluctuations was generated by using Eq. [7] for the 10988 days of climatic data (see end of this section for a description of the model driving data):

$$y(t) = A_s \sin(2\pi f t) \qquad [7]$$

where A_s is the amplitude of the sine curve, which is defined as the maximum displacement of the function from its centre position, and shows the height of the curve (fluctuation) (here, A_s

= 100 cm), 2π is the natural period of the sine curve, f is the frequency (here, $f = 1/365$ days^{-1}), and t is the time period of the sine curve (here, $t = 1$ to 10988 days).

The upper boundary condition in HYDRUS was set to an atmospheric boundary with surface runoff. The domain was non-equally discretized with 401 nodes, with finer discretization at the top to account for the stronger flow dynamics close to the soil surface. Pressure head was used for initialization of the soil profile with linearly decreasing potentials between the bottom (0 cm) and the top (200 cm) node (i.e., hydrostatic equilibrium).

For the simulations, different setups were chosen with varying complexity. A simple homogeneous soil profile without vegetation was selected as the simplest case, to study the impact of the choice of different PTFs. Complexity was increased by adding different vegetation covers (grass and wheat). In both cases, growth was not simulated and both crops covered the soil throughout the entire year. Potential evapotranspiration, ET_0, was split into soil evaporation, E, and transpiration, T, by setting T to 75 % of ET_0. Also, rooting depth was assumed to be the same (0-30 cm) for both vegetation covers with a linear decrease in root density from the top soil layer to the maximum rooting depth. The root water uptake reduction model of Feddes et al. (1978) was used, based on the parameter values of Wesseling (1991) for both grass and wheat vegetation, as taken from the HYDRUS embedded look up table (see Tab. 2). Therefore, the only difference between both vegetation scenarios was the root water uptake. This simplification in terms of growing season and rooting depth was done to simplify the comparison of the simulation results, by ensuring that root water uptake will not be from different soil layers when grass is replaced by wheat. In a next step of increasing complexity, soil layering was introduced, whereby two layering schemes were assumed. 1) Sandy loam over silt loam overlaying a loamy sand and 2) silt loam over silty clay loam overlaying a silty clay, respectively. For the layered profiles the first layer was set to extend from 0 to 50 cm, the second layer from 50 – 100 cm, whereas the third layer occupied the rest of the profile (100 – 200 cm). Again, the same vegetation parameters as for the homogeneous soil were used. Finally, the layered system covered by wheat with a fluctuating groundwater table was simulated. Figure 2 shows a schematic view of the seven model scenarios used in this study.

Thirty years (10988 days) of daily climatic data (comprising precipitation and Penmen-Monteith potential ET) from 1982 to 2011 were taken from North Rhine-Westphalia, Germany (mean NRW climatic data) as used by Hoffmann et al. (2016) and Kuhnert et al. (2017). The climate is humid temperate.

2.4. Statistics and data evaluation

The HYDRUS-1D outputs were used to analyse the fluxes of the soil water balance, i.e. actual evaporation (E_a), actual evapotranspiration (ET_a), drainage (D), and runoff (R). For the data of the 13 different PTFs the cumulative flux was selected and the arithmetic mean, 0.05 and 0.95 percentiles (which spans the 90 percent tolerance interval), and the 0.15 and 0.85 percentiles (which spans the 70 percent tolerance interval), were calculated for the last data point of the cumulative fluxes (t_{end} = day 10988) for each soil textural class. In a next step, outliers based on the percentiles were calculated and flagged according to Eq. [8] and [9]:

$$If\ value\ of\ a\ PTF\ for\ flux\,(x) > 0.95\ percentile,\ value = 1,\ else\ value = 0 \quad [8]$$

$$If\ value\ of\ a\ PTF\ for\ flux\,(x) < 0.05\ percentile,\ value = -1,\ else\ value = 0 \quad [9]$$

where $flux(x)$ is the cumulated flux (e.g., actual evapotranspiration) at t_{end} for each individual PTF used to simulate the soil class for each model scenario. In other words, if the flux value, $flux(x)$, at t_{end} exceeds (is less than) the percentile span calculated, the simulation was flagged with a 1 (-1), whereas if the flux value, $flux(x)$, at t_{end} lies within the given percentile, the simulation was flagged with a 0. The same procedure as for the 90 percent tolerance interval was repeated for the 70 percent tolerance interval. In order to present variability of the hydraulic parameters or simulated fluxes, the coefficient of variation (CV) was calculated, relating the standard deviation to the mean value. The CV was expressed as a percentage.

For the analysis of differences in soil hydraulic properties estimated by the PTFs, the comparison of two group means was performed with the non-parametric Wilcoxon rank sum test using Matlab® *ranksum* function using the probability p = 0.05. Principal component analysis (PCA) was performed, also with Matlab®.

To help interpret the differences between model runs with regards to the water fluxes, the matric flux potential, *MFP*, the characteristic length, L_C, the sorptivity, S, and characteristic time t_{grav} was calculated as explained next.

The matric flux potential *MFP* [cm^2 d^{-1}] is a convenient bulk soil hydraulic property that is often used in soil water movement studies (e.g., Raats, 1977; Pullan, 1990; Grant & Groenevelt, 2015), which is defined as the integral of the hydraulic conductivity $K_{(h)}$ [cm d^{-1}] over the pressure head h [cm] starting at an arbitrary reference pressure head h_{ref}:

$$MFP = \int_{h_{ref}}^{h} K_{(h)}\,dh \quad [10]$$

where h_{ref} was set to permanent wilting point at h = -15,000 cm according to Pinheiro et al. (2018) and h set to full saturation (h = 0).

As a second soil physical feature, the characteristic length of bare soil evaporation, L_C [cm], was calculated. According to Lehmann et al. (2008 and 2018), L_C is the maximal extent of the capillary flow region to supply water to evaporating surface. L_C is determined by the range of capillary pressure between large and small pores driving the capillary flux against gravity (expressed as head difference, denoted as gravity length L_G) and the hydraulic conductivity K_{eff} of the supply region. Formally, L_C is defined via:

$$L_C = \frac{L_G}{1+\frac{E_0}{K_{eff}}} = \frac{\frac{(1-m)}{\alpha}\left(1+\frac{1}{m}\right)^{(1+m)}}{1+\frac{E_0}{4\,K(h_{crit})}} \quad [11]$$

where α and m are the van Genuchten parameters used in Eq. [4] and E_0 is potential evaporation rate. To calculate effective conductivity, K_{eff} was estimated as $4K_{crit}$ (Haghighi et al., 2013; Lehmann et al., 2018), whereby K_{crit} is the unsaturated hydraulic conductivity at critical water content and capillary pressure when capillary pathways start to disconnect (Lehmann et al., 2008). The critical capillary pressure and gravity length are determined based on linearization of the soil water retention curve. For the van Genuchten formulation used in Eq. [11], the linearized retention curve consists of the tangent to the inflection point, and L_G and h_{crit} can be expressed analytically. For Brooks and Corey, the values are determined numerically with a line passing through the air-entry value (S_e = 1, $h = h_b$) and a particular point on the retention curve that is closest to (S_e = 0, $h = h_b$).

The soil sorptivity, S [cm $d^{0.5}$], which is defined as a measure of the capacity of a porous medium to absorb or desorb liquid by capillarity, was calculated assuming a soil column with uniform initial water content and infinite length, following the approach of Parlange (1975) as described in Moret-Fernández et al. (2017), Lotorre et al. (2018), and Rahmati et al. (2019):

$$S^2(\theta_s,\theta_i) = \int_{\theta_i}^{\theta_s} D(\theta)\left[\theta_s+\theta-2\theta_i\right]d\theta \quad [12]$$

where $D(\theta)$ [cm^2 d^{-1}] is the diffusivity defined by Klute (1952) as:

$$D(\theta)=K(\theta)\frac{dh}{d\theta} \quad [13]$$

For the initial water content, θ_i, the maximum reported residual water content (θ_r) for all MvG parameters used in this study (0.192 cm^3 cm^{-3}) was used for all PTFs based on MvG and 0.12 cm^3 cm^{-3} for all PTFs based on BC.

Finally, the so-called characteristic time (t_{grav}) was calculated according to Philip (1957), which determines the 'time' t where gravitational forces become dominant ($t \gg t_{grav}$), while for $t \ll t_{grav}$ capillary forces remain dominant over gravitational forces (Rahmati et al., 2020):

$$t_{grav}=\left(\frac{S}{K_s}\right)^2 \quad [14]$$

Two final related soil properties that were calculated were the characteristic time for the attainment of field capacity *FC*, τ_{FC} (d), and the elapsed time required for attainment of *FC*, denoted as t_{FC} (d]), see Assouline and Or (2014). To do so, the effective soil saturation at field capacity S_{FC} [Eq. 15] and the unsaturated hydraulic conductivity at field capacity $K(S_{FC})$ [Eq. 16] were calculated by:

$$S_{FC}=\left[1+\left\{\left(\frac{n-1}{n}\right)^{(1-2n)}\right\}\right]^{\left(\frac{1-n}{n}\right)} \quad [15]$$

$$K_{FC}=K_s S_{FC}^{0.5} \quad [16]$$

Moreover, the quantity of drainable water from the soil profile to depth z at field capacity *FC*, Q_{FC}, expressed as equivalent water depth (dimensions of length), has to be calculated as:

$$Q_{FC}=z(\theta_s-\theta_r)(1-S_{FC}) \quad [17]$$

Here, z was set to 30 cm to match the maximal rooting depth for convenience, as z only scales with total Q_{FC}.

Consequently, a characteristic time [d] for the attainment of *FC*, τ_{FC}, can be deduced from the ratio of these two quantities, Q_{FC} and K_{FC}:

$$\tau_{FC} = \frac{Q_{FC}}{K_{FC}} \quad [18]$$

The drainage dynamics can now be linked to the elapsed time [d] required for attainment of *FC*, denoted as t_{FC}:

$$t_{FC} = \frac{z(\theta_s - \theta_r)}{K_m} \ln\left(\frac{K(S_{FC})}{Ks}\right) \quad [19]$$

where K_m is the effective hydraulic conductivity that represents the mean value of $K(S_{FC})$ weighted by S_{FC} (representing the available relative cross section of flow):

$$K_m = \frac{\int_0^1 S_{FC} K(S_{FC}) d S_{FC}}{\int_0^1 S_{FC} d S_{FC}} \quad [20]$$

3. Results and Discussion

3.1. Predicted hydraulic parameters and hydraulic functions

In a first step, the retention and hydraulic conductivity curves for the 13 PTFs and the 12 USDA soil classes were plotted based on the estimated soil hydraulic parameters listed in Annex Tab. 1 and 2. As an example, the retention and hydraulic conductivity curves for the USDA textural class sand is plotted in Fig. 3 (see Annex Fig. 1 for all retention and hydraulic conductivity curves of all USDA textural classes). Figure 3 shows that the retention curves based on the 13 PTFs greatly differ along the entire pressure head range. For these curves, the saturated water content, θ_s, varies between 0.472 for Rawls MvG and 0.375 $cm^3\ cm^{-3}$ for Cosby SSC with a mean of 0.417 $cm^3\ cm^{-3}$ over all PTFs. The corresponding coefficient of variation (CV) is 7.5 % for the sandy soil. The residual water content, θ_r, varies between 0 for those PTFs setting θ_r to 0 such as Weynants, Clapp&Hornberger, Cosby SC, and Cosby SSC, to 0.061 $cm^3\ cm^{-3}$ for the 'Toth class' PTF with a mean of 0.029 $cm^3\ cm^{-3}$ and a CV of 80.1%, indicating a much higher variability in terms of CV in θ_r compared to θ_s. An even larger

variability can be found in the saturated hydraulic conductivity, K_s, for which we found a maximum value of 1520.6 cm d^{-1} for Clapp&Hornberger and a minimum of 8.3 cm d^{-1} for the 'Toth class' with a mean of 315.8 cm d^{-1} (CV = 129.2 %). In general, the smallest CV values (data not shown) for all USDA textural classes were found for θ_s, with lowest ranging from 2.4% for the silty clay loam class to CV = 7.5 % for the sand class. (Larger variability was observed in θ_r with the lowest coefficient of variation in the loamy sand class (CV = 76.3 %) and highest in the silty clay loam class (CV = 106.6 %). As expected, K_s showed the largest variability, with lowest CV in the loam class (91.7 %) and largest in the clay loam class (215.2 %). The larger CV values for the K_s estimation is not surprising as a large uncertainty in predicted K_s has been already widely reported (e.g., Jaynes & Tyler, 1984; Ahuja et al., 1985; Tietje & Hennings, 1996; Schaap et al., 1998). Furthermore, it has to be noted that the PTF developed by Weynants et al. (2009) predicts K_s^* instead of K_s, where K_s^* is a hydraulic conductivity acting as a matching point at suction head h = -6. Therefore, some slightly lower K_s (here K_s^*) value will be predicted by Weynants' PTF. On the other hand, there seems to be a clear grouping among the class PTFs, with regards to the estimation of K_s. Clapp&Hornberger predicted the highest values for six classes (sand, loamy sand, sandy loam, loam, silt loam, and silty clay loam), followed by the PTF of Woesten for five soil classes (sandy clay loam, clay loam, sandy loam, silt loam, and clay). Even more pronounced is the picture for the prediction of lowest K_s, whereby the PTF of Rawls MvG estimated the smallest K_s values for 11 soil textural classes, except for sand, whereas the 'Toth class' PTF showed the lowest K_s values. Unfortunately, the α and n (or $1/b$) values cannot be directly compared between the BC and MvG approaches, as both parameters have a slightly different physical meaning.

3.2. Numerical model performance

As numerical stability of the simulation is one of the crucial aspects in the choice and application of the PTF, especially for large scale modelling, we analysed each PTF with respect to numerical convergence, when using HYDRUS. For each PTF and the seven model scenarios (see Fig. 2), 44 individual model runs were performed: for each PTF, the three homogeneous soil layer model scenarios were modelled for each soil textural class (these are 36 model runs). In addition, the four layered configurations are run for a coarse and a fine soil layering, resulting in eight model runs; hence, a total of 44 model runs per PTF were obtained. Note that for the Clapp and Hornberger (1978) PTF, only 41 model runs were performed as no parameters were reported for the silt class. For 486 model runs, out of the

total 569 model runs (i.e., 85 %), convergence was achieved. A total of 184 out of 217 (85%) of the model runs for the BC and 279 out of 352 (79 %) for the MvG parameterization converged, even though it has been reported that the BC type function sometimes prevents rapid convergence and might therefore cause numerical problems. This was deemed to be caused by the discontinuity present in the slope of both the soil water retention and unsaturated hydraulic conductivity curves (van Genuchten, 1980).

For those cases where the simulation did not converge for the MvG parameters, the air entrance value (the inverse of α) was set to -2 cm, and the model was rerun. This procedure increased the total number of converged MvG simulation runs to 302 (86 %), which is a similar percentage to that obtained for BC. Looking at individual PTFs (see Tab. 3), we can see that the Rosetta SSC+BD and Cosby SSC seem to be numerically very stable with 100 % converged runs. The Woesten and 'Toth continuous' function converged for > 95 % of the runs, after setting the $1/\alpha$ to -2 cm. On the other hand, the lowest convergence was found for the Rawls MvG and Rawls BC with 43 and 39 %. Unfortunately, using $1/\alpha$ = -2 cm did not improve convergence for Rawls MvG.

The reason why some PTFs prohibited the HYDRUS model from converging is quite apparent for some cases. For example, Rawls MvG and Rawls BC yielded very low K_s values of 0.8 cm day^{-1} for the loam and sandy clay class and ≤ 0.3 cm day^{-1} for clay loam, silt, and silt loam class. Extremely low values were obtained for silty clay loam (0.04 cm day^{-1}) and clay (0.004 cm day^{-1}), almost allowing no infiltration at all. Another extremely low K_s value was predicted by the 'Toth class' PTF for silty clay, with 0.01 cm day^{-1}; these unrealistically low values again led to numerical instabilities.

It has to be noted that the reported convergence here is only valid for the numerical model (HYDRUS-1D) used in this exercise with the given numerical (convergence) default criteria, vertical discretization and temporal resolution, and atmospheric boundary conditions. The performance of these PTFs may change if a different numerical scheme, e.g. solving the Richards equation in the mixed or diffusivity form were used, or a different spatial discretization and/or temporal resolution. Furthermore, the lack of certain processes in our simulations (e.g. coupled heat and water transport or evaporation from the wet canopy), or the nature of the atmospheric forcings (e.g. a difference in rainfall frequency and amount) will affect the likelihood of convergence.

3.3. Fluxes and outliers

3.3.1 Simulated fluxes over time

Firstly, the simulated cumulative fluxes were analysed. ET_a (vegetated surface) or E_a (bare soil) is a key flux as it indirectly contains information of the net infiltration into the soil profile (net daily infiltration = daily sum of precipitation – daily sum of ET_a or E_a), deep drainage (over long-run), and plant available water in the root zone. Furthermore, ET_a or E_a determines the return of water from the soil profile to the atmosphere, and as such affects the land surface energy budget. Cumulative ET_a or E_a data for each scenario/soil class combination was plotted and the arithmetic mean of all data (model ensemble mean, MEM) for each combination, as well as the spread of the data, was calculated by the 70 and 90 percent tolerance interval according to Eq. [8] and [9].

As an example of the high variability in simulated fluxes, the simulated cumulative E_a, for the homogeneous bare soil scenario of *loamy sand* texture, over the entire simulation period of 30 years, that ends on day 10988 (t_{end}), is plotted in Fig. 4a. There is a large variability between the various simulations based on the 13 PTFs. MEM at t_{end} is 1692 cm (564 mm year^{-1}). The smallest simulated cumulative E_a was 1273 cm (424 mm year^{-1}) for the Carsel&Parrish PTF, and largest, with 2043 cm (681 mm year^{-1}), for Weynants. The difference of 257 mm year^{-1} between the largest and smallest simulated E_a, and their deviation of 140 and 117 mm year^{-1} from MEM clearly indicates that the choice of PTF substantially affects the estimation of the E_a for this soil class. In contrast, low variability was found for cumulative E_a for the bare homogeneous *clay loam* (see Fig. 4b). Notably, two out of the 13 simulations did not converge (Rawls MvG and Rawls BC), which potentially also impacts the variability in simulated fluxes. Nevertheless, for the remaining 11 simulations the lowest simulated flux was 1744 cm (581 mm year^{-1}) for Rawls class PTF and largest for Weynants PTF with 2041 cm (680 mm year^{-1}) (for this soil, MEM = 1893 cm or 631 mm year^{-1}). Overall, the difference between the largest and smallest flux is only 99 mm year^{-1}, i.e., 2.5 times smaller than the difference found for the loamy sand. As E_a will be also be influenced by the precipitation entering the soil (total precipitation over 30 years = 2479.7 cm (827 mm yr^{-1})), we also looked at the cumulative runoff. For most soil textural class/PTF combinations, still for the homogeneous bare soil scenarios, runoff is low or negligible, with zero runoff, or values < 1 cm over 30 years, for 121 model runs, which is equivalent to 88 % of runs. Nine simulations (7 %) returned a runoff >1 cm but <10 cm, and eight exceeded 10 cm over 30 years (7 %). The highest cumulative runoff was generated for the Rawls MvG/silt combination with 675 cm (27 % of total precipitation) followed by Rawls silt loam with 664 cm and Rawls MvG loam with 389 cm. These three combinations have also been flagged as outliers of the lower 0.15 percentile for total cumulative evaporation at t_{end}, which can be explained by the fact that

less water enters the soil, and therefore less will be evaporated. The same holds for the Carsel&Parrish PTF and silty clay loam (runoff = 135.1 cm) and sandy clay (runoff = 70 cm) combinations. On the other hand, the combination 'Toth class' PTF/silt loam generated 122.9 cm runoff but was classified as an upper 0.95 percentile outlier, generating more E_a. Finally, Rawls BC/silt and 'Toth continuous'/silty clay combinations generated runoff of 156.1 and 43.7 cm, respectively, yet are not classified as outliers. In general, runoff generation is linked to low K_s values (see Annex Tab. 1 and 2). An overview of all cumulative E_a fluxes at t_{end} for the bare soil scenarios is plotted in Fig. Annex 2.

Our findings with regards to ET_a for the vegetated scenarios (grass and wheat), still with a homogeneous soil profile (Annex Fig 3 to 4), were comparable. For some soil classes, such as clay, clay loam, and silty clay loam, variability between PTFs was low, for both grass and wheat, whereas the ET_a for the sandy and sandy loam soils showed consistently high variability. In contrast, ET_a for the loamy sand class exhibited relatively high variability for grass (as was also the case for bare soil) and a slightly smaller one for the wheat scenario configuration. For the other soil textural classes, the picture is less clear. Again, as was the case for the bare soil, there is a substantial number of soil class/PTF combinations that result in runoff. A slightly larger, compared to the bare soil scenario, percentage of simulations with runoff > 1 cm was found for the grass (18 %) and the wheat (20 %) scenarios. Moreover, maximum runoff at t_{end} value increased from bare (675 cm for Rawls MvG silt) via grass (859.2 cm for Rawls MvG silty clay) to the wheat scenario (999.2 cm for Rawls BC silty clay). Surprisingly, eight out of twelve soil class/PTF combinations yielding runoff > 100 cm were not flagged as outliers for the 90 % tolerance interval for the grass and four out of 10 for the wheat. There are some unexpected findings, namely that the simulation for the Toth continuous' PTF yielded 46.1 and 11.7 cm runoff, respectively, for the silty clay and silty clay loam under wheat vegetation, despite the fact that the ET_a flux at t_{end} was flagged as an outlier of the upper 0.95 percentile, indicating relatively high evaporation with respect to the model ensemble.

Finally, the simulation for the scenario of sandy loam overlying silt loam and loamy sand plotted in Annex Fig 5 showed much lower variability for E_a compared to E_a of the homogenous profile with the texture of the uppermost layer (silt loam in Annex Fig. 2). This indicates that soil layering will reduce the effect of the choice of PTF on the cumulative evaporation. This holds true even more for the layered bare soil scenario where silt loam overlies silty clay loam that is overlying silty clay. Again, the variability in E_a for the layered system is much lower than that of the homogeneous silt loam, that forms the first layer in the

vertically heterogeneous soil profile. Besides, it can clearly be seen that when vegetation is introduced, variability increases slightly, which is reflected in the coefficient of variation (CV) of the flux at t_{end}, where for the layered profile topped by sandy loam the CV increased from 3.9, via 5.1 to 4.9 % for the bare, grass, and wheat vegetation scenario. For the profile with the first layer consisting of silt loam, CV values were 0.5, 3.7, and 3.7 % for the bare, grass, and wheat vegetation, respectively. Introducing a fluctuating ground water table increased the CV substantially to 13.6 % for both vegetated layered systems. Variability in simulated E_a or ET_a for the layered scenarios can partly be explained by a large reduction in runoff. In total only two simulations (2 %) for the Carsel&Parrish (sandy loam topped layered profile under grass vegetation (290.2 cm) and sandy loam topped layered profile for the wheat vegetation and ground water fluctuation (302.6 cm)) exceeded runoff of 100 cm. A further four exceeded the runoff threshold of 1 cm (3 combinations for Carsel&Parrish and one for Rawls BC).

Overall, the choice of PTF substantially affects the simulated values of E_a or ET_a for most soil classes, irrespective of the fact whether the soil was bare, where the water (vapour) can only leave the soil column via the pore-space at the soil surface, or vegetated, where a considerable proportion of the water being returned to the atmosphere consist of water taken up from the deeper rooted parts of the soil profile.

3.3.2. Outliers per scenario

As shown above, substantial variability in simulated E_a or ET_a fluxes occurred for different PTFs and model scenarios. The fluxes exceeding the 70 or 90 % tolerance intervals, respectively, were marked as outliers and calculated for each scenario and soil class according to Eq. [8] and [9]. The number of outliers were counted for each scenario individually in a first step.

The number of outliers varies greatly between PTFs with regards to E_a fluxes for the homogeneous bare soil (see Fig. 5a); these fluxes were shown in Figs. 4a and b. Naturally, more outliers are detected for the 70 than for the 90 % tolerance interval. For this scenario, Rosetta SSC, Weynants, and 'Toth class' exceed the upper 0.95 percentile, whereby Weynants exceeded this percentile for all soil classes where the model had converged, except for silt and silt loam. Rosetta SSC exceeded the upper 0.95 percentile for clay, whereas the 'Toth class' PTF exceeded it for silt and silt loam, respectively. Looking at the lower 5 percentile, Carsel&Parrish PTF exceeded this threshold for eight soil classes, and further outliers were found for Rawls MvG (N = 6) and Rawls BC class (N = 4). Two outliers were calculated for Cosby SSC, Rawls BC, and one for Rosetta SSC and Woesten PTFs.

Finally, only Rosetta SSC+BD, 'Toth continuous', and Cosby SC indicate no outliers for the upper and lower 70 % tolerance interval.

As some simulation runs did not converge (see discussion above), the comparison in terms of total number of outliers is limited. Therefore, the total number of outliers was normalized to the number of converged simulations for each scenario and PTF combination. Again, we present the relative number of outliers for the homogeneous bare soil profile simulations in Fig. 5b as an example (all others are shown in Annex Fig. 6 to 8). Here, the Weynants PTF shows the largest percentage of outliers for the upper 0.95 and 0.85 percentile with 82 and 100 % outliers, respectively. For the 'Toth class' PTF, we found 20 and 50 % outliers for the upper 0.95 and 0.85 percentile, respectively, indicating that also this PTF simulated larger fluxes with respect to the ensemble. On the other hand, Rawls MvG shows the largest percentage of outliers at the lower end (86 % for the 0.15 and 57 % for the 0.05 percentile) followed by Carsel&Parrish PTF with 73 % for the 0.15 and 45 % for the 0.05 percentile. However, Rawls BC and Rawls class also show substantial percentages of outliers for the 0.15 percentile. By comparing the relative (converged only) and absolute (all runs) number of outliers, it can be seen that despite equal or even lower or higher absolute number of outliers for different PTFs, the relative numbers differ due to non-converged simulation runs for some PTFs. For instance, 'Toth class' for the 0.95 percentile showed 2 outliers yielding 20 % relative outliers as two simulations (silty clay and silty clay loam) did not converge, whereby 1 outlier for Rosetta SSC yielded only 8 % relative outliers as all simulations converged.

As there is no clear trend in the analysis of the absolute or relative outliers for the individual scenarios (see Fig 5b and Annex Figs. 6b to 8b) which PTF generates most outliers, from here on the outliers over all scenarios for all soil textural classes were calculated for converged simulation runs only and expressed in relative terms. Figure 6a shows the outliers of the 90 % tolerance interval (sum of upper and lower outliers), combined for all textural classes, for the seven scenarios for the 13 PTFs for E_a and ET_a at t_{end}. In this figure, the PTFs of the two main hydraulic formulations are clustered: those based on the Mualem van Genuchten (MvG) on the left and those based on Brooks Corey (BC) formulation on the right. Furthermore, two lines are added, dividing the results into three groups: i) those PTFs with relative number of outliers < 10 %, classified as 'robust', ii) those PTFs with 10 % ≥outliers ≤ 20 %, classified as 'intermediate robust', and iii) the PTFs with relative number of outliers >20 %, classified as 'non-robust'. It has to be noted that these thresholds (10 and 20 %) were chosen arbitrarily, but may help to formulate the final recommendations for the choice of preferred PTF, to be used in land surface models, for example.

This classification shows that the Rosetta SSC, Rosetta SSC+BD, Woesten, 'Toth continuous', Rawls BC, Rawls class BC, Clapp&Hornberger, Cosby SC, and Cosby SSC PTFs are located below the 10 % threshold for the 90 % tolerance interval, and can be therefore classified as 'robust' with respect to the ensemble behaviour (spread). Interestingly, all PTFs using BC formulation show low relative numbers of outliers below 10%. Woesten PTF did not show any outliers at all, indicating that this PTF is very robust with respect to the PTF ensemble used. On the other hand, the 'Toth class' PTF was classified as intermediate robust, and three PTFs (Carsel&Parrish, Rawls MvG, and Weynants) were classified as non-robust, whereby Rawls MvG produced most outliers (32 %).

The results for the 70 % tolerance interval are shown in Fig. 6b and followed the same approach as for the 90 % tolerance interval discussed above. Four PTFs are characterised as robust (Rawls BC, Clapp&Hornberger, Cosby SC, and Cosby SSC). Again, all these four PTFs serve to produce parameters for the Brooks Corey hydraulic formulation. The intermediate robust grouping includes Rosetta SSC, Rosetta SSC+BC, Woesten, and 'Toth continuous', that provide parameters for the Mualem van Genuchten formulation. Finally, Carsel&Parrish, Rawls MvG, Weynants, 'Toth class', and Rawls BC class are those PTFs classified as non-robust. There are two class, rather than continuous, PTFs here, indicating that continuous PTFs are more likely to be robust. Also, the Weynants PTF was based on a relative small number of samples, for Belgium only.

Based on the results presented above, it can be concluded that the use of different PTFs results in different hydraulic properties that predict considerably different E_a or ET_a fluxes leading to different soil water contents in the root zone but also to differences in deep percolation (or ground water recharge). Furthermore, PTFs such as Carsel&Parrish, Rawls MvG and Weynants can be identified as systematically less robust. In contrast, others, such as Woesten or all PTFs using the Brooks Corey formulation (except Rawls BC class) seem to be robust with respect to the ensemble of PTFs used in this study.

To facilitate the identification of outliers, all outliers per PTF, scenario, and textural class combination were colour coded and plotted in Tab. 4. Again, Weynants overestimates E_a or ET_a fluxes (brown colour for dryer soil conditions) for nearly all textural soil classes except for clay, and silt. On the other hand, Rawls MvG shows underestimation (blue colour for wetter soil conditions) for loam and silt loam over all three homogeneous soil scenarios and for silt and sandy loam for two out of the three homogeneous soil scenarios. The Carsel&Parrish' PTF, on the other hand, results in over- and underestimation, depending on soil class.

In the study of Zheng et al. (2020), who evaluated also a set of 13 PTFs using independent retention data (data not used in the development of the PTFs), the Carsel&Parrish PTF showed largest RMSE in predicted volumetric water content of the retention data points, which is in agreement to our findings that those PTF is characterized as less robust with respect to simulated fluxes. On the other hand, Weynants PTF was the one with lowest RMSE, whereas Weynants was characterized as less robust in our case. The reason why Weynants behaves differently between the study presented by Zheng et al. (2020) and our study, might be that only the retention characteristics were analysed by Zheng et al. (2020), whereas in the function sensitivity performed here, also the hydraulic conductivity function plays a virtual role. The reasons why Weynants is characterized as less robust is further discussed in the following sections.

3.3.3. Simulated spread with respect to scenario

We raised the hypothesis that differences (variability) in simulated fluxes from using different PTFs will be reduced with increasing model complexity. Increasing complexity was generated by introducing vegetation (grass or wheat), soil layering, or the assumption of a fluctuating ground water table, for the layered vegetated soil scenario only. As only the homogeneous scenarios (bare, grass, and wheat) used all soil classes, we restrict the analysis on these three scenarios.

For the analysis, again the simulated cumulative actual E_a or ET_a data at t_{end} was taken and the model ensemble mean (MEM) for E_a to ET_a at t_{end} over all PTFs was calculated for each individual soil class and scenario. Based on the MEM value for E_a *or* ET_a, as well as the individual E_a or ET_a value at t_{end} for each model run, the % difference from the MEM ($100/\text{MEM}*E_a@t_{end}$ or $ET_a@t_{end}$) was calculated and visualized using boxplots in Fig. 7, where the red line indicates the median, the box indicates the 0.25 and 0.75 percentiles, the whiskers represent the most extreme data points not considered as outliers, and crosses represent the outliers (value is more than 1.5 times the interquartile range). From the boxplots, two types of information can be deduced: i), the variability of predicted E_a or ET_a over all PTFs for one soil class / scenario and ii), the change in variability (spread) resulting from a change in scenario complexity (bare, grass, or wheat vegetation).

In general, the largest variability in predicted E_a or ET_a was found for the bare soil conditions, which is most pronounced for the loam, loamy sand, sand, sandy clay, clay loam, and sandy loam class. Minor differences were found between bare and vegetated scenarios for the other soil classes. The silty clay soil class for the grass scenario showed the smallest overall spread

between minimum and maximum predicted E_a (or ET_a) with a value of 7 % (min = 97 % and max 104 %). On the other hand, the largest variability was found for the combination sandy soil/bare soil scenario with 53 % (min = 72 % and max 125 %). All spreads, throughout the 13 PTFs, for different soil classes and scenarios are provided in the final column of Tab. 4. Overall, bare scenarios show a mean spread of 30 %, whereby the grass and wheat vegetated scenarios have only 23 % spread over all soil classes. A possible explanation for the reduced spread in simulated E_a or ET_a with increasing model complexity (in this case vegetation) is that for the vegetated profiles water is extracted from the rooted portion of the soil profile, whereas under bare soil the water can only leave the soil profile at the soil surface. In the latter case, differences in the soil hydraulic properties, especially in unsaturated hydraulic conductivity, which is highly variable (on the order of magnitudes) between PTFs, close to the surface will impact the E_a flux more substantially. As shown earlier, runoff will occur in both scenarios (bare and vegetated) and is even slightly larger for the vegetated scenario, and therefore, cannot explain the reduced variability.

Next, for the layered soil scenarios (bare, grass, and wheat, without fluctuating groundwater table) a clear reduction in the variability was observed, for both profiles (by sandy loam or silt loam). The bare and the vegetated scenarios showed nearly the same spread (mean 13.4 % for bare, 18.5 % for grass, and 15.5 % for the wheat). In general, the sandy loam overlaying silt loam and loamy sand showed always higher variability compared to the silt loam overlaying silty clay loam and silty clay, which is consistent to the finding that the sandy loam of the homogeneous soil profile also showed higher variability compared to the homogeneous silt loam scenarios.

Overall, the results indicate that adding vegetation reduces the variability in the simulated E_a or ET_a flux, even if runoff occurs more frequently. This conclusion also holds for adding more complexity in terms of soil layering, although the latter has to be regarded with some caution due to the low number of soil combinations selected for these model runs. However, taking into account that large portions of our global land surface is covered by vegetation, differences in predicted fluxes, as a result of differences in PTFs used to generate the hydraulic parameters, will most likely be smaller compared to an ‘unvegetated world’.

In contrast, adding a fluctuating ground water table to the layered wheat scenario greatly increased variability in ET_a flux, for both soil layering to 48 and 35 % for the sandy loam overlaying silt loam and loamy sand, and silt loam overlaying silty clay loam and silty clay, respectively.

3.3.4. Differences in instantaneous fluxes

Cumulative fluxes at t_{end} will only provide long-term systematic under or overestimation, but will not provide information on how the instantaneous fluxes fluctuate compared to the MEM. Therefore, the instantaneous fluxes were also analysed. The same analyses as conducted for the cumulative fluxes were performed, i.e., calculation of the MEM and the 0.95 and 0.05 percentiles for time step i, whereby i runs from day 1 to 10988. Secondly, the total number as well as the upper and lower percentile outliers were counted. As an example, the outliers of E_a for the sandy loam of the homogeneous bare soil scenario were plotted in Fig. 8, for the different PTFs. Carsel&Parrish PTF shows a substantial number of outliers for the lower 0.05 percentile (N = 2020 or 18 % of all days), indicating that for these days less water will evaporate and return to the atmosphere, which would have implications for the cloud forming processes of a numerical weather prediction or climate model if a LSM using this PTF were to be embedded within it. On the other hand, Weynants has 3053 outliers for the upper 0.85 percentile (28 %) but also a smaller number of outliers for the lower 0.05 percentile (N = 309 or 3 %), leading to larger E_a flux. A large number of 0.05 percentile outliers were also found for Rawls MvG (N = 2992 or 27 %), again combined with a lower number of upper 0.95 percentile outliers (N = 391 or 4 %). Cosby SC, Cosby SSC, and Rawls BC showed only low number of outliers (N < 10) for the upper and lower percentiles. Even though the model runs for which the hydraulic parameters were derived from Carsel&Parrish and Rawls MvG PTFs exhibit large numbers of outliers, both are not flagged as 90 % tolerance interval outliers when the cumulative flux at t_{end} was analysed. This means that the non-flagged instantaneous E_a fluxes compensate for the lower fluxes determined as outliers in Fig 8, or that the outliers are close to the 0.15 percentile, which is reflected by the fact that the total sum of underestimated flux (outlier flux – flux for the lower 0.05 percentile for each outlier day) is low, amounting to 5.7 and 2.4 cm over the 30-year period, respectively. Moreover, both PTFs show runoff exceeding a total of 1 cm in 60% (Carsel&Parrish) and 36 % (Rawls MvG) of all converged simulations, respectively. For the Rawls MvG the nine simulations with runoff even exceed the 100 cm threshold, with runoff ranging between 388.6 to 859.2 cm. Looking at all textural classes (data not shown) for the homogeneous bare soil scenario, 29 soil class / PTF combinations out of the total 151 do not exhibit any outliers at all for the instantaneous E_a flux. These outliers are clustered in three soil classes only (clay, silty clay, and silty clay loam). Interestingly, out of these 29 with zero outliers in instantaneous flux, five are flagged as outliers for the cumulative flux at t_{end} (Rosetta SSC clay, Cosby SSC clay, Weynants silty clay and silty clay loam, as well as

Carsel&Parrish silty clay loam), meaning that these PTFs over- or underestimate instantaneous E_a only very modestly, yet consistently throughout the simulation period.

The percentage of all 90 percent tolerance outliers (sum of upper and lower outliers) summed over all days for all three homogeneous soil scenarios (bare, grass, and wheat) for all soil classes and PTFs are provided in Tab. 4. Over all soil classes and PTFs, the bare soil scenario has the lowest total number of outliers (N = 119930 days or 6.5 % over all days and scenarios) followed by the homogeneous wheat configuration (N = 173961 days or 10.1 %) and the homogeneous grass scenario (N = 178249 days or 10.4 %). This finding is perhaps in contradiction to the finding that the percental spread in cumulative E_a or ET_a at t_{end} was larger for the bare soil scenario, compared to the vegetated ones. Furthermore, for some texture classes the total number of outliers increased remarkably when vegetation was implemented, such as for the clay class, where the bare soil scenario has no outliers (0%), while the percentage of outliers increased to 14 % for the homogeneous grass and wheat scenario, respectively. This indicates that the differences in available root zone water, affecting actual transpiration, are the main driver for differences between PTFs, compared to fluxes over the soil surface E_a. On the other hand, only the silty clay and the silty clay loam showed no outliers at all for the instantaneous flux for all scenarios. Looking at all soil class/PTF/scenarios combinations, no clear trend in the total number of outliers in instantaneous evapo(transpi)ration flux, and flagged outliers for the cumulative E_a or ET_a flux at t_{end} can be observed. This leads to the conclusion that the outliers in instantaneous flux alone do not necessarily sum up to a cumulative flux flagged as an outlier.

3.4.Explaining variability and outliers by soil physical properties

As has been shown, substantial variability exists in cumulative and instantaneous fluxes, and some PTFs are found to be more robust than others. In this section, we discuss in more detail the reasons for the differences between the predicted soil water fluxes, resulting from the use of different PTFs, by analysing the estimated hydraulic parameters K_s, λ (MvG tortuosity parameter) and the soil physical characteristics. In general, variability between estimated K_s for the different PTFs is quite low (Fig. 9a), and values for Rawls MvG and BC only are significantly lower than all other PTFs. These lower values may explain the poor numerical convergence for these simulations, and the prevalence of lower E_a fluxes as well as a high number of lower 0.05 percentile outliers at t_{end} as depicted in Tab. 4, especially for Rawls MvG. Clapp&Hornberger K_s values are significantly higher than those estimated by the Weynants PTF, ‘Toth class’, and Cosby SC and SSC, yet did not show any high outliers for E_a fluxes. Interestingly, Cosby SC and SSC were developed based on the same water retention

and K_s data as Clapp&Hornberger, as both used data from Holtan et al. (1968), nevertheless estimated K_s values are quite different. One reason might be that Clapp&Hornberger only used textural classes, and averaged K_s for those classes, whereas Cosby SC and SSC is a continuous PTF. Coming back to the outliers listed in Tab. 4, those runs based on Weynants PTF indicate larger E_a fluxes at t_{end} and a large number of upper 0.95 percentile outliers, whereas their estimated K_s is not significantly different from most other PTFs. Here, it has to be noted that Weynants did not estimate K_s but rather estimated a near saturation hydraulic conductivity K_s* that is mainly controlled by textural properties and which is lower that K_s. The results suggest that variability in K_s alone cannot explain the flux differences simulated.

Looking at the λ value used in MvG formulation, two different classes of PTFs can be distinguished, those setting λ to 0.5 as originally proposed by van Genuchten (1980) (Carsel&Parrish, Rawls MvG, and 'Toth continuous') and those who fitted λ as an additional free parameter (Rosetta SC and SSC, Woesten, Weynants, and 'Toth class'). The variability in λ is plotted in Fig. 9b. It shows that the λ estimates of Weynants' PTFs are significantly lower than those from the other four PTFs estimating λ, except for 'Toth class'. 'Toth class' λ values are significantly lower than those calculated by Rosetta SC and SSC, and than those setting λ to 0.5, whereas Woesten is significantly lower than Rosetta SC and SSC, and < 0.5. The more negative λ values for Weynants appear strongly related to the larger number of upper 0.95 percentile outliers listed in Tab. 4, whereas the intermediately low λ values for 'Toth class' and Woesten PTF do not explain the number of flagged outliers. In general, λ is significantly correlated to the MvG parameter n for those PTFs setting $\lambda \neq 0.5$ (R^2 =0.40, p = 0.05, data not shown) indicating a nonlinear behaviour which can be described as $n = 1.58\,e^{0.064\,\lambda}$ with an R^2 of 0.51. Looking at the ranges of λ, there is a systematic difference between PTFs, with largest λ values for Rosetta (-3.1>λ<0.62), followed by Woesten ((-4.46>λ<0.60), 'Toth class' (-5.5>λ<0.73), and Weynants (-7.87>λ<1.92). Rosetta and Woesten are characterized by low numbers of tolerance interval outliers, whereas 'Toth class' and Weynants are characterized by large number of tolerance outliers, both in the upper end (upper 0.95 percentile outliers). As λ is correlated to the n parameter, and n directly impacts the hydraulic properties and hence L_G, L_C, τ_{FC}, and t_{FC}, and to a less extend S, the correlation between λ and these soil characteristics was calculated. The results indicated (data not shown) that λ is not significantly correlated to L_C, t_{grav}, τ_{FC}, and t_{FC} but moderately correlated to L_G (R^2 =0.31, p = 0.05) and S (R^2 =0.30, p = 0.05), whereas λ is not correlated to the flux E_a at t_{end}.

For the calculated soil characteristics L_G, Weynants shows large variability and high median and significantly differs from Woesten, Rawls MvG, Rawls BC class, Rawls BC and Cosby SC and SSC. In contrast, Rawls BC and BC class show low L_G, and Rawls BC class is significant different from Rosetta SSC and Clapp&Hornberger (see Fig. 10a). Here, it has to be kept in mind that L_G solely depends on the water retention characteristics and hence the n and α values play a crucial role in the calculation. As n is positively correlated with λ, and Weynants shows the smallest λ values, the significant difference, with regards to L_G, between Weynants and most other PTFs seems logical. Large L_G values occur for very fine textures, which are classically associated to low K values that limit water supply to the evaporating surface, which is reflected by the higher number of upper 0.95 percentile outliers for Weynants, leading to a drier soil profile. Lower E_a fluxes at t_{end}, and therefore, a wetter profile occurred frequently for Carsel&Parrish and Rawls (MVG and BC), whereby all these PTFs are also located in the low L_G range.

The calculation of L_C is based on knowledge of L_G and the actual hydraulic conductivity distribution above the evaporation front. Therefore, K_s plays also an important role in the calculation of L_C. The impact of K_s on L_C is clearly reflected in the high L_C values for Clapp & Hornberger, which exhibit high K_s values across all soil classes compared to all other PTFs (see Fig. 10b). At the other end of the spectrum, the impact of K_s on L_C is also apparent for Rawls MvG and Rawls BC which do not indicate much spread and are characterized by low K_s and hence low L_C. Surprisingly, Clapp&Hornberger are not classified as outliers when looking at cumulative fluxes (see Tab. 4), whereas the low L_C for Rawls MvG corresponds to the number of outliers detected. On the other hand, Weynants, which was characterized as the PTF with most outliers at the upper 0.95 percentile, lies in the middle of the range of L_C values depicted in Fig. 10b, indicating that L_C might not be a good indicator for flagged outliers. As stated in Lehmann et al. (2008), L_C longer than 1 m are considered as unrealistic (evaporative extraction of water by capillary flow across several meters is unlikely). Interestingly, only the Clapp&Hornberger PTF show $L_C > 1$ m, while all other PTFs give realistic values.

The analysis of *MFP* shows a quite different picture (Fig. 10c). Here, the PTFs based on Brooks Corey group together and exhibit a higher *MFP* compared to the MvG based PTFs. Testing on significance showed that Rawls BC class, Clapp&Hornberger, and both Cosby PTFs are significantly different from all others and that only Rawls BC is not significantly different from those using MvG formulation, except for Rawls MvG. This is of interest, as Rawls MvG is only a ‘translation’ of the Brooks Corey to van Genuchten parameters from

Rawls BC according to Morel-Seytoux (1986), while keeping K_s. As the Weynants PTF showed substantial outliers, as listed in Tab. 4, one would also expect Weynants to be different with regards to *MFP* as the λ value is much smaller compared to all other PTFs, while K_s does not differ (see Fig. 9a and 9b). One reason for the fact that *MFP* for Weynants does not differ from the other PTFs might be its relatively low *n* value, as λ and *n* are positively correlated. The impact of λ as opposed to the effect of MFP becomes clearer when we compare Weynants and Woesten, which show no significant difference in *MFP*, yet larger K_s values for Woesten and lower λ for Weynants. Overall, the *MFP* cannot explain the outliers detected and depicted in Tab. 4 as only Rawls MvG is systematically different and exhibits large number of outliers, whereas Weynants *MFP* are in the centre of the range of values found for the different PTFs. On the other hand, *MFP* values for Clapp&Hornberger, as well as for both PTFs from Cosby, are significantly higher, yet do not stand out in Tab. 4.

With regards to the sorptivity *S* (Fig. 11a), there is a large variability in *S* for Clapp&Hornberger, which is significantly different from all other PTFs. Small variabilities in *S*, however, are found for Woesten, Rawls MvG and BC, Weynants, 'Toth class' and both Cosby PTF. In general, *S* is moderately correlated to L_C ($R^2 = 0.40$).

Rawls BC shows a high t_{grav}, which is significantly different from all other PTFs, except for Rawls MvG. Both Cosby PTFs and both Rawls continuous functions (Rawls MvG and BC) show relatively large variability (Fig. 11b). The higher t_{grav} for Rawls MvG fits with the larger number of outliers listed in Tab. 4, whereas for BC this pattern is not clear, maybe due to the lack of numerical convergence. In general, larger t_{grav} values are associated with more fine-grained soils such as loam and clays (Alastal, 2012), whereas the low t_{grav} of Woesten characterizes more coarse soils such as sands.

High τ_{FC} were calculated for Rawls MvG (Fig. 11c), whereby the large τ_{FC} is associated with extremely low predicted K_s values. Extremely high values were found for Rawls MvG with τ_{FC} exceeding 3 million days, whereby Rawls MvG has K_s values of 0.01 and 0.004 cm d^{-1} for the silty clay and clay class, respectively, and also did not converge. For the two soil classes, silt and silt loam, where the model run did converge τ_{FC} is also extremely large (>44,000 days) and for these soils again low K_s values of 0.2 and 0.3 cm d^{-1}, respectively, were estimated. Additionally, these two model runs are also outliers at the lower 0.05 percentile. Clapp&Hornberger PTF resulted in the smallest τ_{FC}, whereas the K_s predictions are in general higher as for the other soils (see Fig. 9a) and none of the simulations were flagged as outliers. On the other hand, all other PTFs have comparable τ_{FC} values, and the outliers detected in Tab. 4 seem not to be linked with τ_{FC}.

Finally, t_{FC} was analysed, which shows the same pattern as τ_{FC}, which is to be expected as t_{FC} and τ_{FC} are linearly correlated as also shown by Assouline and Or (2014).

As these soil physical characteristics were calculated to help explain differences in simulated E_a at t_{end}, all characteristics were correlated against E_a at t_{end} (see Fig. 12). Only $\log_{10}(L_G)$ shows a moderate correlation to E_a at t_{end} (R^2 = 0.52, p =0.05) and a weak correlation was found for $\log_{10}(L_C)$, with R^2 = 0.29 (p =0.05). As E_a, and also drainage D at t_{end}, will be biased if runoff is generated (because less water will infiltrate into the soil profile and be available for evaporation and drainage), E_a and drainage at t_{end} were normalized (E_{a_norm}, D_{norm}) by dividing E_a or drainage at t_{end} by the difference of precipitation at t_{end} (2479.72 cm) and runoff at t_{end}. By doing so, the correlation between λ and E_{a_norm} increased to R^2 =0.31 (p = 0.05). For the derived soil characteristics the correlation also increased (to R^2 = 0.57; p =0.05) for $\log_{10}(L_G)$ but decreased for $\log_{10}(L_C)$, to R^2 = 0.10. On the other hand, the correlation slightly increased for t_{FC}, from R^2 = 0.09 to 0.22.

In a next step, a principal component analysis (PCA) using all converged model runs and soil hydraulic parameters available for MvG and BC (θ_r, θ_s, K_s) as well as all soil characteristics (L_c, L_G, MFP, S, and t_{grav}, t_{FC}, and τ_{FC}) and fluxes (E_{a_tend}, E_{a_norm}, D_{tend}, and D_{norm}) was performed on log transformed data (except θ_r, θ_s, MFP, E_{a_norm}, and D_{norm}) and the results are plotted in Fig. 13. The first three components explain 76 % of the variability in the data and the important loadings on PC 1 (42.9 % of variability) are t_{FC} (0.38), K_s (-0.35), and L_G (0.33). PC 2 (24.5 % of variability) includes the important loadings L_C (0.47), E_a at t_{end} (0.40), and S (0.31). PC 3 explains only 8.6 % of the variability and t_{grav} (0.48) and θ_s (0.47) are the important loadings. The PCA triplot shows scatter of the individual PTFs around the origin of the triplot but also distinct PTF clusters, whereby Weynants (black circle) is oriented along the PC 1 in a fairly small volume and is positively correlated to t_{FC} and τ_{FC} and negatively to D at t_{end} (as drainage D at t_{end} is negative per definition). Rawls (MvG and BC) is oriented in the same direction as Weynants but it exhibits larger scatter, whereas Clapp&Hornberger (red solid markers) is oriented along PC 2 and correlates positively with K_s. K_s values reported by Clapp&Hornberger are amongst the highest compared to all other PTFs as already discussed in relation to Fig. 9a.

Out of these 13 PTFs, three (Clapp&Hornberger, Weynants, and Rawls) can be identified as being distinctive from all others in the triplot as they do not cluster around the origin. Furthermore, they do not only differ considerably in their estimated soil hydraulic parameters (e.g., λ and n value for Weynants, and K_s for Rawls and Clapp&Hornberger) but also in the soil characteristics derived from these parameters, whereby in all soil characteristics either the

n value (remember that *n* is correlated to λ) as well as K_s are directly or indirectly integrated. For example, the low L_C values for Rawls PTFs indicate that the maximum extent of the flow region sustaining evaporation is much smaller than for all other PTFs. This results in low E_a at t_{end} compared to other PFTs and larger number of outliers as depicted in Tab. 4.

Finally, a multiple regression was performed to test whether E_a at t_{end} can be predicted by the soil hydraulic parameters and/or characteristics, whereby only one of those parameters or characteristics were used in turn, i.e. those that were available for MvG and BC. As per Fig. 13, all entries were log transformed except for θ_r, θ_s, and *MFP*, and the best regression was selected using bootstrapping. The best predictive model was found by E_a @ t_{end} = 1252.13 + 183.30 $\log_{10}(L_G)$ + 367.88 $\log_{10}(L_C)$ - 405.22 $\log_{10}(S)$ with an R^2 of 0.88 (see Fig. 14) pointing to the fact that the soil characteristics L_G, L_C, and S describe well the physical behaviour of soils with regards to actual evaporation. Using E_{a_norm} instead of E_a decreased the predictive power of the multiple regression (R^2 = 0.75).

4. Summary and Conclusion

In this study 13 pedotransfer functions (PTF) were used to populate the hydraulic parameters required in the HYDRUS model that was then used to simulate the water fluxes for 12 USDA soil classes, for different model scenarios that varied in complexity (homogeneous or layered soil profile, with and without vegetation) over a period of 30 years. Plotting the hydraulic functions (water retention and hydraulic conductivity curves) for all PTFs revealed large differences, especially for the hydraulic conductivity curve, leading to the hypothesis that the different PTFs will also show substantial differences in simulated fluxes.

It turned out that some PTFs generated parameters that rendered the HYDRUS model numerically unstable, so that it failed to converge for certain soil class/configuration combinations, especially those reported by Rawls and Brakensiek (1985) (Rawls MvG) and by Rawls et al. (1982) (Rawls BC), which converged only in less than 44 % off all simulation runs. Surprisingly, PTFs using the Brooks Corey (BC) formulation resulted in higher convergence rates, compared to those based on Mualem van Genuchten, even though BC is in general perceived to be less stable.

In a next step, differences in simulated actual evaporation E_a or evapotranspiration ET_a between the model runs were analysed, as E_a and ET_a indirectly contain information on the net infiltration, deep drainage (over long-term) and water stored in the root zone. Therefore, the cumulative E_a or ET_a at the end of the simulation period (t_{end} = 10988 days) was selected and the 90 and 70 % tolerance interval as well as the model ensemble mean were calculated.

Fluxes exceeding the tolerance limits were flagged and counted. The results indicate that some PTFs (Rawls MvG, Weynants, and Carsel&Parrish) were classified as non-robust, as the fluxes generated by the parameters derived from these PTFs exceeded a defined threshold of 20 % of the 90 % tolerance interval outliers over all scenarios and soil classes. On the other hand, all PTFs using the Brooks Corey formulation (Rawls BC, Rawls BC class, Clapp&Hornberger, Cosby SC, and Cosby SSC) are classified as robust, as they generally result in a low percentage of 90 % tolerance outliers. The PTF of Woesten performed best, and it showed no outliers at all for the 90 % tolerance interval. A hypothesis raised at the beginning of the study was that increasing model complexity will reduce the variability in predicted fluxes. Therefore, the individual simulated E_a and ET_a fluxes at t_{end} were compared to the model ensemble mean (MEM), and the relative spread of the individual simulations was calculated. The results show that the bare soil scenarios exhibit the highest mean percentage spread (30 %), whereas the grass and wheat vegetated scenarios had a reduced spread (23%), averaged over all soil classes. The reduction in relative spread with the inclusion of vegetation can be explained by the fact that for these runs the water leaving the soils can be extracted from the entire rooted soil profile (after which it gets transpired via the vegetation), whereas under bare soil conditions it can only leave the soil profile at the soil surface. In the latter case, differences in the soil hydraulic properties close to the surface, especially in unsaturated hydraulic conductivity, which is highly variable (in order of magnitudes) between PTFs, will impact the E_a or ET_a flux more substantially.

The instantaneous E_a or ET_a fluxes over time were also analysed, whereby again the 90 % tolerance outliers were calculated and counted. The results indicate that some PTF/soil class/ model scenario combinations showed substantial outliers in the instantaneous fluxes, yet were not flagged as outliers for the cumulative flux at t_{end}, indicating that the non-flagged instantaneous fluxes compensate these outliers. On the other hand, other PTF/soil class/scenario combinations showed no outliers for the instantaneous fluxes, but were flagged as outliers for the cumulative case, indicating that even small over- or underestimations in instantaneous flux can sum up to large errors in the long-run.

To explain differences in simulated E_a for the homogeneous bare soil scenario, different soil characteristics were calculated, and a PCA was conducted using all simulated fluxes, soil hydraulic parameters and soil characteristics available for both MvG and BC. The PCA revealed three distinct PTFs clusters, namely Weynants, Rawls, and Clapp&Hornberger, whereby Weynants and Rawls were also characterized by a large number of tolerance outliers. Weynants correlates positively to gravity time of infiltration t_{FC} and τ_{FC} and negatively to

drainage D at t_{end}, whereby Clapp&Hornberger is oriented in the opposite direction and correlated with the saturated conductivity K_s. For Rawls a reasonable correlation with t_{FC} and τ_{FC} is found, but due to the large scatter for this PTF the interpretation is less clear.

Finally, a multiple regression was performed, showing, that the gravitational length L_G, characteristic length of evaporation L_C and sorptivity S together explain almost 90% of the variability in simulated E_a at t_{end}.

Overall, our results provide insights in the functional behaviour of the PTFs as a bases for the selection of PTFs in land surface modelling, but also for large scale hydrological or crop models, where considerations regarding the numerical stability, model behaviour and performance over the long run and instantaneously should be balanced against each other. Based on this, Rosetta SSC+BD, Woesten, and 'Toth continuous' seem to be the most robust PTFs for the Mualem van Genuchten function and Cosby SC for Brooks Corey. Note, however, that our study is in essence a sensitivity analysis; it does not include model verification using measured fluxes, and it employs one model only.

In any case, the results clearly demonstrate that the choice of PTF can substantially affect the simulated fluxes, and as a consequence, the water content stored in the soil profile with part of that available for root water uptake and crop growth. Therefore, we strongly recommend to harmonize the PTFs used in land surface, large scale hydrological, or crop model inter-comparison studies to avoid artefacts originating from the choice of PTF rather than from model structures. Additionally, our study should motivate future studies, where measured verification fluxes are available from lysimeters and or eddy covariance stations.

Acknowledgements

The authors would like to acknowledge Yakov Pachepsky and Attila Nemes for the help with Rawls PTF.

Data Availability Statement

The HYDRUS-1D code is freely available online (https://www.pc-progress.com/en/Default.aspx?h1d-downloads). The atmospheric data are also freely available online (https://doi.org/10.7910/DVN/C0J5BB).

References

Ahuja, L., Naney, J.W., & Williams, R.D. (1985). Estimating Soil Water Characteristics from Simpler Properties or Limited Data. *Soil Science Society of American Journal*, 49 (5), 1100-1105. doi:10.2136/sssaj1985.03615995004900050005x.

Alastal, K. (2012). Oscillatory flows and capillary effects in partially saturated and unsaturated porous media: applications to beach hydrodynamics. (Ecoulements oscillatoires et effets capillaires en milieux poreux partiellement saturés et non satu-rés: applications en hydrodynamique côtière). Doctoral thesis of Institut National Polytech. de Toulouse / niv. de Toulouse. Institut de Mécanique des Fluides de Toulouse. Toulouse, France, 226 pp.

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.

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

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

Carsel, R.F. & Parrish, R.S. (1988). Developing joint probability distributions of soil water retention characteristics. *Water Resources Research*, 24. 755-769.

Chirico, G.B., Medina, H. & Romano, N. (2010). Functional evaluation of PTF prediction uncertainty: An application at hillslope scale. *Geoderma*, 155. 193–202.

Christiaens, K. & Feyen, J. (2001). Analysis of uncertainties associated with different methods to determine soil hydraulic properties and their propagation in the distributed hydrological MIKE SHE model. *Journal of Hydrology*, 246. 63–81.

Clapp, R.B. & Hornberger, G.M. (1978). Empirical Equations for some soil hydraulic properties. *Water Resources Research*, 14 (4). 601-604.

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. 638–648.

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. 683-690.

Donatelli, M., Wösten, J.H.M. & Belocchi, G. (2004). Evaluation of pedotransfer functions. In: Y. Pachepsky and W.J. Rawls (Eds.) Development of Pedotransfer functions in soil hydrology. Elsevier, Amsterdam. p. 357–362.

Ek, M.B. & Holtslag, A.A.M. (2004). Influence of Soil Moisture on Boundary Layer Cloud Development. *Journal of Hydrometeorology*, 5, 86–99. https://doi.org/10.1175/1525-7541(2004)005<0086:IOSMOB>2.0.CO;2

Espino, A., Mallants, D., Vanclooster, M. & Feyen, J. (1995). Cautionary notes on the use of pedotransfer functions for estimating soil hydraulic properties. *Agricultural Water Management Journal*, 29. 235-253.

Feddes, R.A., Bresler, E., & Neuman, S.P. (1974). Field test of a modified numerical model for water uptake by root systems. *Water Resources Research*, 10. 1199-1206.

Grant, C.D. & Groenevelt, P.H. (2015). Weighting the differential water capacity to account for declining hydraulic conductivity in a drying coarse-textured soil. *Soil Research*. 53, 386–391. doi:10.1071/SR14258.

Groh, J., Vanderborght, J., Pütz, T., Vogel, H.J., Gründling, R., Rupp, H., Rahmati, M., Sommer, M., Vereecken, H., & Gerke, H.H. (2020). Responses of soil water storage and crop water use efficiency to changing climatic conditions: A lysimeter-based space-for-time approach. *Hydrology and Earth System Sciences*, 24 (3). https://doi.org/10.5194/hess-2019-411.

Guber, A.K., Pachepsky, Y.A., van Genuchten, M.Th., Rawls, W.J., Simunek, J., Jacques, D., Nicholson, T.J. & Cady, R.E. (2006). Field-Scale Water Flow Simulations Using Ensembles of Pedotransfer Functions for Soil Water Retention. *Vadose Zone Journal*, 5(1). 234-247. doi: 10.2136/vzj2005.0111.

Haghighi, E., Shahraeeni, E., Lehmann, P., & Or, D. (2013). Evaporation rates across a convective air boundary layer are dominated by diffusion. *Water Resources Research*, 49(3). 1602-1610.

Hoffmann, H., Zhao, G., Asseng, S., Bindi, M., Biernath, C., Constantin, J., Coucheney, E. Dechow, R., Doro, L., Eckersten, H., Gaiser, T., Grosz, B., Heinlein, F., Kassie, B.T., Kersebaum, K.-C., Klein, C., Kuhnert, M., Lewan, E., Moriondo, M., Nendel, C., Priesack, E., Raynal, H., Roggero, P.P., Rötter, R.P., Siebert, S., Specka, X. , Tao, F. Teixeira, E., Trombi, G., Wallach, D., Weihermüller, L., Yeluripati, J., & Ewert, F. (2016). Impact of Spatial Soil and Climate Input Data Aggregation on Regional Yield Simulations. *PLOS ONE.* doi:10.1371/journal.pone.0151782.

Holtan, H.N., England, C.B., Lawless, G.P., & Schumaker, G.A. (1968). Moisture-tension data for selected soils on experimental water-sheds, Rep. ARS 41-144. Agric. Res. Serv., Beltsville, Md. 609 pp.

Iwema, J., Rosolem, R., Rahman, M., Blyth, E., & Wagener, T. (2017). Land surface model performance using cosmic-ray and point-scale soil moisture measurements for calibration, *Hydrology and Earth System Science*, 21. 2843–2861. https://doi.org/10.5194/hess-21-2843-2017.

Jaynes, D.B. & Tyler, E.J. (1984). Using soil physical properties to estimate hydraulic conductivity. *Soil Science*, 138. 298-305.

Kuhnert, M, Yeluripati, J., Smith, P., Hoffmann, H., van Oijen, M., Constantin, J., Coucheney, E., Dechow, R., Eckersten, H., Gaiser, T., Grosz, B., Haas, E., Kersebaum, K.-C., Kiese, R., Klatt, S., Lewan, E., Nendel, C., Raynal, H., Sosa, C., Specka, X., Teixeira, E., Wang, E., Weihermüller, L., Zhao, G., Zhao, Z., Ogle, S., & Ewert, F. (2017). Impact analysis of climate data aggregation at different spatial scales on simulated net primary productivity for croplands. *European Journal of Agronomy*, 88. 41-52. doi: EURAGR-25545.

KAK. (1994). Bodenkundliche Kartieranleitung. 4th edition. E. Schweizerbart'sche Verlagsbuchhandlung. Stuttgart.

Koster, R.D. & Suarez, M.J. (2001). Soil Moisture Memory in Climate Models. *Journal of Hydrometeorolgy*, 2. 558–570. https://doi.org/10.1175/1525-7541(2001)002<0558:SMMICM>.

Klute, A.J. (1952). Some theoretical aspects of the flow of water in unsaturated soils. *Soil Science of America Proceedings*, 16. 144–148.

Latorre, B., Moret-Fernández, D., Lassabatere, L., Rahmati, M., López, M.V., Angulo-Jaramillo, R., Sorando, R., Comín, F. & Jiménez, J.J. (2018). Influence of the β parameter of the Haverkamp model on the transient soil water infiltration curve. *Journal of Hydrology*, 564. 222-229.

Lehmann, P., Assouline, S. & Or, D. (2008). Characteristic lengths affecting evaporative drying of porous media. *Physical Review E*, 77(5). 056309. doi:10.1103/PhysRevE.77.056309.

Lehmann, P., Merlin, O., Gentine, P., & Or, D. (2018). Soil texture effects on surface resistance to bare soil evaporation. *Geophysical Research Letters*, 45. 10,398–10,405. https://doi.org/10.1029/2018GL078803.

Morel-Seytoux, H.J., Meyer, P.D., Nachabe, M., Touma, J., van Genuchten, M.T., & Lenhard, R.J. (1986). Parameter equivalence for the Brooks-Corey and van Genuchten soil characteristics: Preserving the effective capillary drive. *Water Resources Research*, 32(5). 1251-1258.

Moret-Fernández, D., Latorre, & Angulo-Martínez, M. (2017). Comparison of different methods to estimate the soil sorptivity from an upward infiltration curve. *Catena*, 155. 86-92.

Neyshaboury M.R., Rahmati, M., Rafiee Alavi, S.A.R., Rezaee, H., & Nazemi, A. H. (2015). Prediction of unsaturated soil hydraulic conductivity by using air permeability: Regression approach. *Indian Journal of Agricultural Research*, 49 (6). 528-533.

Parlange, J.Y. (1975). On solving the flow equation in unsaturated flow by optimization: horizontal infiltration. *Soil Science Society of America Journal*, 39. 415–418.

Philip, J.R. (1957). The theory of infiltration: 1. The infiltration equation and its solution. *Soil Science*, 83. 345-358.

Pinheiro, E.A.R., de Jong van Lier, Q., & Metselaar, K. (2018). A matric flux potential approach to assess plant water availability in two climate zones in Brazil. *Vadose Zone Journal*, 17.1600 83. doi:10.2136/vzj2016.09.0 083.

Pullan, A.J. (1990). The quasilinear approach for unsaturated porous media flow. *Water Resources Research*, 26.1219–1234. doi:10.1029/WR026i006p01219.

Raats, P.A.C. (1977). Laterally confined, steady flows of water from sources and to sinks in unsaturated soils. *Soil Science Society of America Journal*, 41. 294–304. doi:10.2136/sssaj1977.03615995004100020025x.

Rahmati, M., Groh, J., Graf, A., Pütz, T., Vanderborght, J., and Vereecken, H. (2020). On the impact of increasing drought on the relationship between soil water content and evapotranspiration of a grassland. *Vadose Zone Journal*, 19. e20029. https://doi.org/10.1002/vzj2.20029.

Rahmati M., Latorre, B., Lassabatere, L., Angulo-Jaramillo, R., Moret-Fernández, D. (2019). The relevance of Philip theory to Haverkamp quasi-exact implicit analytical formulation and its uses to predict soil hydraulic properties. *Journal of Hydrology*, 570.: 816–826.

Rahmati, M., & Neyshaboury, M.R. (2016). Soil Air Permeability Modeling and Its Use for Predicting Unsaturated Soil Hydraulic Conductivity. *Soil Science Society of America Journal*, 80(6). 1507-1513.

Rahmati M., Neyshabouri, M.R., Doussan, C., & Behroozinezhad, B. (2013). Simplified estimation of unsaturated soil hydraulic conductivity using bulk electrical conductivity and particle size distribution. *Soil Research*, 51. 23–33.

Rahmati M., Vanderborght, J., Šimunek, J., Vrugt, J.A., Moret-Fernández, D., Latorre, B., Lassabatere, L., & Vereecken, H. (2020). Soil hydraulic properties estimation from one-dimensional infiltration experiments using characteristic time concept. *Vadose Zone Journal*, https://doi.org/10.1002/vzj2.20068 .

Rawls, W.J. & Brakensiek, D.L. (1985). Prediction of soil water properties for hydrologic modeling. p 293-299. In: Jones, E.B. and T.J. Ward (eds.). Watershed Management in the Eighties. Proc. Irrig. Drain. Div., ASCE, Denver, CO. April 30 - May 1, 1985., 34: 3293-3302.

Rawls, W.J., Brakensiek, D.L., & Saxton, K.E. (1982). Estimating soil water properties. *Transactions ASAE*, 25(5). 1316-1320 and 1328.

Schaap, M.G. (2004). Accuracy and uncertainty in PTF predictions. p. 33–43. In: Y. Pachepsky and W.J. Rawls (Eds.) Development of pedotransfer functions in soil hydrology. Elsevier, Amsterdam.

Schaap, M.G., & Leij, F.J. (1998). Database-related accuracy and uncertainty of pedotransfer functions. *Soil Science*, 163. 765–779.

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. 163–176.

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. 27-34. doi:10.2136/vzj2005.0005.

Shao, Y. & Irannejad, P. (1999). On the Choice of Soil Hydraulic Models in Land-Surface Schemes. *Boundary-Layer Meteorology*, 90. 83–115. https://doi.org/10.1023/A:1001786023282

Šimůnek, J. & van Genuchten, M.T. (2008). Modelling nonequilibrium flow and transport processes using HYDRUS. *Vadose Zone Journal*, 7. 782-797.

Šimůnek, J., van Genuchten, M. T., & Šejna, M. (2008). Development and applications of the HYDRUS and STANMOD software packages and related codes. *Vadose Zone Journal*, 7(2). 587-600.

Seneviratne, S.I., Corti, T. Davin, E.L., Hirschi, M., Jaeger, E.B., Lehner, I., Orlowsky, B., & Teuling, A.J. (2010). Investigating soil moisture–climate interactions in a changing climate: a review. *Earth Science Reviews*, 99 (3–4). 125–161.

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.

Tietje, O. & Tapkenhinrichs, M. (1993). Evaluation of pedo-transfer functions. *Soil Science Society of America Journal*, 57 (4). 1088-1095.

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

van den Hurk, B., Ettema, J., & Viterbo, P. (2008). Analysis of Soil Moisture Changes in Europe during a Single Growing Season in a New ECMWF Soil Moisture Assimilation System. *Journal of Hydrometeorolgy*, 9. 116–131. https://doi.org/10.1175/2007JHM848.1.

van Genuchten, M.Th. (1980). A closed form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, 44. 892 898.

van Genuchten, M.T. & Pachepsky, Y. (2011). Pedotransfer functions. In: J. Glinski, J. Horabik, and J. Lipiec (Eds.). Encyclopedia of Agrophysics. Springer: 556-561. doi: 10.1007/978-90-481-3585-1_109.

van Looy K., Bouma, J., Herbst, M., Koestel, J., Minasny, B., Mishra, U., Montzka, C., Nemes, A., Pachepsky, Y., Padarian, J., Schaap, M., Tóth, B., Verhoef, A., Vanderborght, J., van der Ploeg, M., Weihermüller, L., Zacharias, S., Zhang, Y. & Vereecken, H. (2017). Pedotransfer functions in Earth system science: challenges and perspectives. *Reviews of Geophysics*. DOI:10.1002/2017RG000581.

Vereecken, H., Diels, J. Vanorshoven, 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. 1371–1378.

Vereecken, H., Huisman, J.A., Bogena, H., Vanderborght, J., Vrugt, J.A., & Hopmans, J.W. (2008). On the value of soil moisture measurements in vadose zone hydrology: a review. *Water Resources Research*, 44. W00D06, doi:10.1029/2008WR006829. 21.

Vereecken, H., Huisman, J.A., Pachepsky, Y., Montzka, C., van der Kruk, J., Bogena, H., Weihermüller, L., Herbst, M., Martinez, G., & Vanderborght, J. (2014). On the spatio-

temporal dynamics of soil moisture at the field scale. *Journal of Hydrology*, 516, 76–96.

Vereecken, H., Maes, J., & Feyen, J. (1990). Estimating unsaturated hydraulic conductivity from easily measured soil properties. *Soil Science*, 149. 1–12.

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. 389–403.

Vereecken, H., Weynants, M., Javaux, M., Pachepsky, Y., Schaap, M.G., & van Genuchten, M.T. (2010). Using pedotransfer functions to estimate the Van Genuchten-Mualem soil hydraulic properties- a review. *Vadose Zone Journal*, 9. 795–820. doi:10.2136/vzj2010.0045

Vereecken, H., Weihermüller, L., Assouline, S., Šimůnek, J., Verhoef, A., Herbst, M., Archer, N., Mohanty, B., Montzka, C., Vanderborght, J., Balsamo, G., Bechtold, M., Boone, A., Chadburn, S., Cuntz, M., Decharme, B., Ducharne, A., Ek, M., Garrigues, S., Goergen, K., Ingwersen, J., Kollet, S., Lawrence, D.M., Li, Q., Or, D., Swenson, S., de Vrese, P., Walko, R., Wu, Y., & Xue, Y. (2019). Infiltration from the Pedon to Global Grid Scales: An Overview and Outlook for Land Surface Modeling. *Vadose Zone Journal*, 18(1). doi: 10.2136/vzj2018.10.0191.

Vogel, T., van Genuchten, M.Th., & Cislerova, M. (2001). Effect of the shape of the soil hydraulic functions near saturation on variably-saturated flow predictions. *Advances in Water Resources*, 24 (2). 133-144 https://doi.org/10.1016/S0309-1708(00)00037-3.

Wagner, B., Tarnawski, V.R., Hennings, V., Müller, U., Wessolek, G., & Plagge, R. (2001). Evaluation of pedo-transfer functions for unsaturated soil hydraulic conductivity using an independent data set. *Geoderma*, 102. 275–297.

Weihermüller, L., Herbst, M., Javaux, M., & Weynants, M. (2017). Erratum to "Revisiting Vereecken Pedotransfer Functions: Introducing a Closed-Form Hydraulic Model. *Vadose Zone Journal*, doi:10.2136/vzj2008.0062er

Wesseling, J.G., Eibers, J.A., Kabat, P., & van den Broek, B.J. (1991). SWATRE, Instructions for input. Internal note, Win and Staring Centre, Wageningen.

Weynants, M., Vereecken, H., & Javaux, M. (2009). Revisiting Vereecken Pedotransfer Functions: Introducing a Closed-Form Hydraulic Model. *Vadose Zone Journal*, 8 (1). 86-95. doi: 10.2136/vzj2008.0062.

Wösten J.H.M., Lilly, A., Nemes, A., & Le Bas, C. (1999). Development and use of a database of hydraulic properties of European soils. *Geoderma*, 90.169-185.

Wösten, J.H.M., Pachepsky, Y.A., & Rawls, W.J. (2001). Pedotransfer functions: Bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*, 251. 123–150.

Yakirevich, A., Pachepsky, Y.A., Gish, T.J., Guber, A.K., Kuznetsov, M.Y., Cady, R.E., & Nicholson, T. (2013). Augmentation of groundwater monitoring networks using information theory and ensemble modeling with pedotransfer functions. *Journal of Hydrology*, 501. 13-24. doi: 10.1016/j.jhydrol.2013.07.032.

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.

Zhang, Y., Schaap, M.G., Wei, Z. (2020). Development of hierarchical ensemble model and estimates of soil water retention with global coverage. *Geophysical Research Letters,* 47, e2020GL088819.

Figure Captions

Figure 1: Overview of the used soil textures for the 12 USDA soil classes. Red points indicate mean texture for each soil class.

Figure 2: Schematic sketch of the seven different scenarios used for the modelling study with increasing model complexity from left to right.

Figure 3: Retention (left) and hydraulic conductivity curves right) for the for the USDA sand class for the 13 PTF (Parameters listed in Annex Tab. 1 and 2). Note that y-axis for the hydraulic conductivity is in log-scale.

Figure 4: a) simulated cumulative actual evaporation E_a [cm] over the simulation period of 10988 days (30 years) for a bare soil with a homogeneous loamy sand soil texture. b) and a bare soil with a homogeneous clay loam soil texture. Light and dark grey shaded area represent the 70 and 90 % tolerance interval, respectively

Figure 5: Absolute and b) relative number of outliers for simulated E_a or ET_a at t_{end} (10988 days) for the 13 pedotransfer functions over all 12 USDA soil classes (11 for Clapp&Hornberger) and the homogeneous bare soil scenario.

Figure 6: a) Absolute and b) relative number of outliers for simulated E_a or ET_a at t_{end} (10988 days) for the 13 pedotransfer functions over all USDA soil classes and scenario depict in Fig. 2. Dotted lines represent a threshold of 10 and 20 % outliers, separating robust (<10 %), intermediate (10-20%), and non-robust (>20%) pedotransfer functions.

Figure 7: Boxplots of relative difference in % from model ensemble mean (MEM) of simulated actual evapotranspiration, ET_a, at t_{end} = 10988 for the homogeneous soil profiles either with bare soil (E_a only) or vegetated with grass or wheat. Red line indicates the median, box the 25 and 75 percentile, whiskers the most extreme data points not considered as outliers, and stars the outliers.

Figure 8: Actual evaporation [cm day^{-1}] for the sandy loam for the homogeneous bare soil scenario and all 12 pedotransfer functions with outliers exceeding the 90 % tolerance interval.

Figure 9: a) boxplots of $\log_{10} K_s$ for all PTFs, b) λ for the MvG formulation, with indication of significant differences. For significance: A differs from all other PTFs if no A is indicated, B, C, and D differ between single PTFs). Boxes are the same as for Fig. 7.

Figure 10: Boxplots for a) L_G, b) L_C, and c) $\log_{10} MFP$, for all PTFs. Boxes and indication of significant differences are the same as for Fig. 7.

Figure 11: Boxplots for a) S b) $\log_{10} t_{grav}$ for all PTFs and c) $\log_{10} \tau_{FC}$ for all PTFs. Boxes and indication of significant differences are the same as for Fig. 7.

Figure 12: Scatterplots of the different soil characteristics gravitational length L_G, characteristic length of evaporation L_C, matrix flux potential MFP, sorptivity S, characteristic time t_{grav}, elapsed time for the attainment of field capacity t_{FC}, characteristic time for the attainment of field capacity τ_{FC}, versus E_a at t_{end} for the homogeneous bare soil scenario as well as τ_{FC} versus t_{FC}.

Figure 13: Trioplot of the principle component analysis for soil parameters, soil characteristics, and fluxes both available for MvG and BC. Note, that only the combination of soil parameters, soil characteristics and converged model runs were used.

Figure 14: Predicted E_a at t_{end} [cm] by multiple regression of soil characteristics $\log_{10}(L_G)$, $\log_{10}(L_C)$, and $\log_{10}(S)$ versus simulated E_a at t_{end} [cm].

Table Captions

Table 1: Overview of used pedotransfer functions (PTFs) with input parameters. MvG = Mualem- van Genuchten type (Eq. 4-5), BC = Brooks-Corey type (Eq. 1-3 Class = class transfer functions of USDA soil classes. Continuous = based on discrete soil information (see PTF Inputs), BD = bulk density (g cm^{-3}), pH (here set to pH = 6.5), and cation exchange capacity CEC (here set to CEC = 0.5*clay+0.005*silt according to KAK (1994)).

Table 2: Feddes root water uptake parameters for pasture (here used for the grass scenario) and wheat vegetation according to Wesseling (1991).

Table 3: Overview of converged simulations for the different PTFs listed in Tab.1. Last two columns indicate the use of air entrance value of -2 cm for the Mualem- van Genuchten type function if simulation did not converge using predicted PTF parameters.

Table 4: Overview of 90% tolerance interval outliers per textural class and scenario (homogeneous bare, grass, or wheat) for E_a/ET_a at t_{end} = 10988 days and the spread over all PTFs from the model ensemble mean (MEM) (colour coded in blue and brown). Numbers for the individual pedotransfer / soil class / scenario combinations depict the % of total 90 % tolerance interval outliers for the instantaneous E_a/ET_a flux. NaN are non-converged simulations. % spread is the spread in % between minimum and maximum cumulative E_a/ET_a at t_{end} over one soil class / scenario combination.

Table 1: Overview of used pedotransfer functions (PTFs) with input parameters. MvG = Mualem- van Genuchten type (Eq. 4-5), BC = Brooks-Corey type (Eq. 1-3), Class = class transfer functions of USDA soil classes. Continuous = based on discrete soil information (see PTF Inputs), BD = bulk density (g cm^{-3}), pH (here set to pH = 6.5), and cation exchange capacity CEC (here set to CEC = 0.5*clay+0.005*silt according to KAK (1994)).

Pedotransfer Function	*General Information*				*PTF Inputs*						*Region*	*No samples*	*Disturbed samples*	$\theta_{(h)}$	$K_{(h)}$	mixed
	MvG	*BC*	*Class*	*Continuous*	*Texture*	*BD*	C_{org}	*Porosity*	*pH*	*CEC*						
Carsel&Parrish (Carsel and Parrish, 1988)	x		x								USA	2942	No	x		
Rosetta SSC (Schaap et al., 2001)	x			x	x						USA / Europe	2134	No	x	x	x[1]
Rosetta SSC+BD (Schaap et al., 2001)	x			x	x	x					USA / Europe	2134	No	x	x	x[1]
Woesten (Wösten et al., 1999)	x			x	x	x	x				Europe	4030	No	x	x	x[1]
Rawls MvG (Rawls and Brakensiek, 1985)	x			x	x			x*			USA	5320	No	x		
Rawls BC (Rawls and Brakensiek, 1985)		x		x	x			x*			USA	5320	No	x		
Rawls BC class (Rawls et al., 1982)		x	x								USA	5320	No	x		
Weynants (Weynants et al., 2009)#	x			x	x	x	x				Belgium	166	No	x	x	x[2] (82%)
Toth continous (Tóth et al., 2015; topsoil)	x			x	x	x	x		x		Europe	4749	No	x	x	x[1] (34%)
Toth class (Tóth et al., 2015; Annex Tab.19)	x		x							x	Europe	-	No	x	x	x[1]
Clapp&Hornberger (Clapp and Hornberger, 1978)		x	x								USA	1446	No	x		
Cosby SC (Cosby et al., 1984)		x		x	x						USA	1446	No	x		
Cosby SSC (Cosby et al., 1984)		x		x	x						USA	1446	No	x		

* porosity was calculated from bulk density by *Porosity* = 1-(*BD*/2.62) SC = sand and clay, SSC = sand, silt, and clay, SSC+BD = sand, silt, caly, and bulk density

Weynants et al. (2009) in combination with Weihermüller et al. (2017)

[1] = fitted on retention data ($\theta_{(h)}$) only and [2] jointly fitted on retention (θ(h)) and hydraulic conductivity (K(h)) data

Table 2: Feddes root water uptake parameters for pasture (here used for the grass scenario) and wheat vegetation according to Wesseling (1991).

Parameter	***Pasture***	***Wheat***
P0 [cm]	-10	0
$P0_{opt}$ [cm]	-25	-1
P2H [cm]	-200	-500
P2L [cm]	-800	-900
P3 [cm]	-8000	-16000
r2H [cm d^{-1}]	0.5	0.5
r2L [cm d^{-1}]	0.1	0.1

Table 3: Overview of converged simulations for the different PTFs listed in Tab.1. Last two columns indicate the use of air entrance value of -2 cm for the Mualem- van Genuchten type function if simulation did not converge using predicted PTF parameters. Note that total number of simulations for Clapp&Hornberger is only 41 as no data for the silt class are reported.

Pedotransfer Function	*General Information*		*converged*		α = -2 cm	
	MvG	*BC*	*n*	*%*	*N*	*%*
Carsel&Parrish (Carsel and Parrish, 1988)	x		30	68	8	86
Rosetta SSC (Schaap et al., 2001)	x		42	95	-	-
Rosetta SSC+BD (Schaap et al., 2001)	x		44	100	-	-
Woesten (Wösten et al., 1999)	x		42	95	1	98
Rawls MvG (Rawls and Brakensiek, 1985)	x		19	43	-	-
Rawls BC (Rawls and Brakensiek, 1985)		x	17	39	-	-
Rawls BC class (Rawls et al., 1982)		x	43	98	-	-
Weynants (Weynants et al., 2009)	x		28	64	12	91
Toth continous (Tóth et al., 2015; topsoil)	x		40	91	2	95
Toth class (Tóth et al., 2015; Annex Tab.19)	x		34	77	-	-
Clapp&Hornberger (Clapp and Hornberger, 1978)		x	37	90	-	-
Cosby SC (Cosby et al., 1984)		x	43	98	-	-
Cosby SSC (Cosby et al., 1984)		x	44	100	-	-

Table 4: Overview of 90% tolerance interval outliers per textural class and scenario (homogeneous bare, grass, or wheat) for E_a/ET_a at t_{end} = 10988 days and the spread over all PTFs from the model ensemble mean (MEM) (colour coded in blue and brown). Numbers for the individual pedotransfer / soil class / scenario combinations depict the % of total 90 % tolerance interval outliers for the instantaneous E_a/ET_a flux. NaN are non-converged simulations. % spread is the spread from the MEM in % between minimum and maximum cumulative E_a/ET_a at t_{end} over one soil class / scenario combination.

scenario	soil class	Carsel&Parrish	Rosetta SSC	Rosetta SSC+BD	Woesten	Rawls MvG	Rawls BC	Rawls BC class	Weynants	Toth continuous	Toth class	Clapp&Hornberger	Cosby SC	Cosby SSC	% spread
homogeneous bare	clay	0	0	0	0	NaN	NaN	0	NaN	0	0	0	0	0	12
homogeneous grass		35	13	18	16	NaN	NaN	4	42	17	19	6	1	10	11
homogeneous wheat		39	15	19	12	NaN	NaN	6	37	23	12	7	2	7	9
homogeneous bare	Clay loam	26	7	7	7	NaN	NaN	6	28	7	6	16	0	0	16
homogeneous grass		0	0	0	0	NaN	NaN	0	0	0	0	NaN	0	0	12
homogeneous wheat		0	0	0	0	NaN	NaN	0	0	0	0	NaN	0	0	13
homogeneous bare	Loam	1	6	5	5	43	NaN	1	29	7	7	10	0	0	30
homogeneous grass		27	6	11	7	72	NaN	4	16	8	18	16	0	2	32
homogeneous wheat		19	10	13	7	54	NaN	3	18	10	21	26	0	1	22
homogeneous bare	Loamy sand	50	2	2	3	9	0	0	31	4	9	6	0	0	46
homogeneous grass		31	5	18	5	68	5	2	15	7	18	8	0	0	40
homogeneous wheat		33	6	19	6	41	5	3	18	8	23	10	0	0	21
homogeneous bare	Sand	56	2	2	3	7	0	0	33	3	5	10	0	0	53
homogeneous grass		28	19	9	5	62	4	3	20	7	12	10	0	0	38
homogeneous wheat		31	19	9	7	36	4	3	21	9	15	16	0	0	39
homogeneous bare	Sandy clay	37	7	7	7	NaN	NaN	0	34	9	8	1	0	1	26
homogeneous grass		24	13	12	10	NaN	NaN	2	55	24	25	7	1	8	11
homogeneous wheat		15	14	12	14	NaN	NaN	2	51	28	29	7	1	8	12
homogeneous bare	Sandy clay loam	0	6	6	5	45	1	0	31	6	6	6	0	0	39
homogeneous grass		29	12	10	13	NaN	27	6	30	11	17	25	1	1	18
homogeneous wheat		20	13	12	15	NaN	28	5	32	13	12	28	0	1	18
homogeneous bare	Sandy loam	18	3	3	3	31	0	1	31	6	7	9	0	0	35
homogeneous grass		24	6	5	6	33	22	1	25	8	7	15	0	0	35
homogeneous wheat		NaN	6	6	8	45	22	2	24	10	8	16	1	1	40
homogeneous bare	Silt	0	7	6	7	43	0	1	6	7	30	NaN	0	4	39
homogeneous grass		10	7	18	8	77	NaN	12	16	8	29	NaN	1	3	34
homogeneous wheat		9	11	19	9	72	NaN	8	15	8	29	NaN	1	5	30
homogeneous bare	Silt loam	0	6	5	5	43	NaN	1	6	6	22	21	0	1	29
homogeneous grass		15	5	13	13	71	NaN	6	19	9	14	20	1	3	32
homogeneous wheat		12	5	16	10	64	NaN	4	14	8	18	33	1	2	26
homogeneous bare	Silty clay	NaN	0	0	0	NaN	NaN	0	0	0	NaN	0	0	0	13
homogeneous grass		NaN	0	0	0	NaN	NaN	0	0	0	NaN	0	0	0	7
homogeneous wheat		NaN	0	0	0	NaN	0	0	0	0	NaN	0	0	0	32
homogeneous bare	Silty clay loam	0	0	0	0	NaN	NaN	0	0	0	NaN	0	0	0	20
homogeneous grass		0	0	0	0	NaN	NaN	0	0	0	NaN	0	0	0	9
homogeneous wheat		0	0	0	0	NaN	NaN	0	0	0	0	0	0	0	8

outlier upper 0.95 percentile

outlier lower 0.05 percentile

class < 10%

class 10-20 %

class 20-30 %

class > 30 %

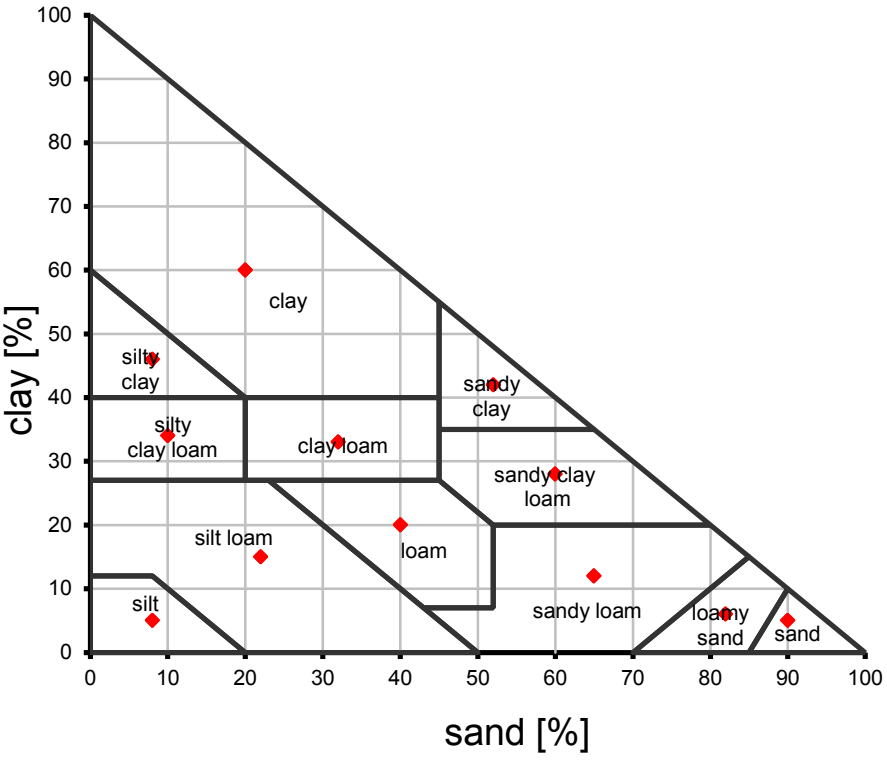


Figure 1: Overview of the used soil textures for the 12 USDA soil classes. Red points indicate mean texture for each soil class.

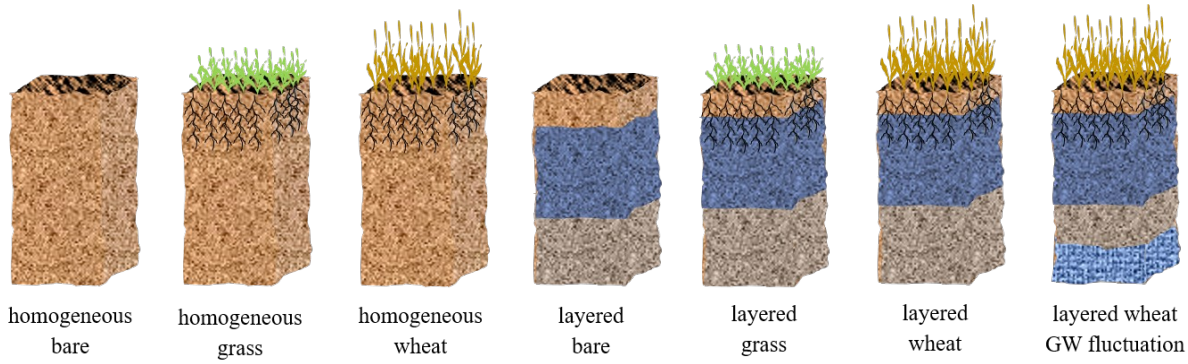
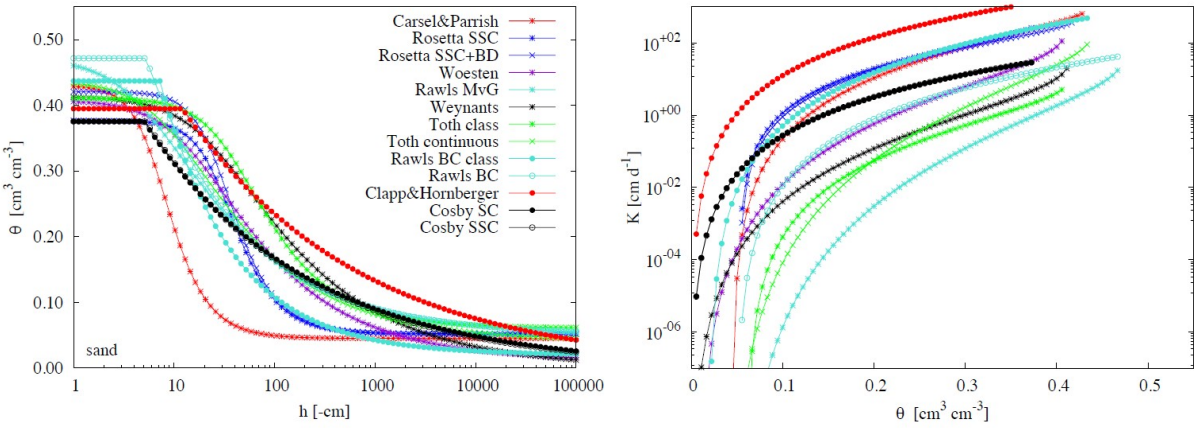


Figure 2: Schematic sketch of the seven different scenarios used for the modelling study with increasing model complexity from left to right. The first three scenarios were computed for each soil textural class. The four layered scenarios were run for two different types of layering.



1390 **Figure 3:** Retention (left) and hydraulic conductivity curves right) for the USDA sand class for the 13 PTF (Parameters listed in Annex Tab. 1 and 2). Note that y-axis for the hydraulic conductivity is in log-scale.

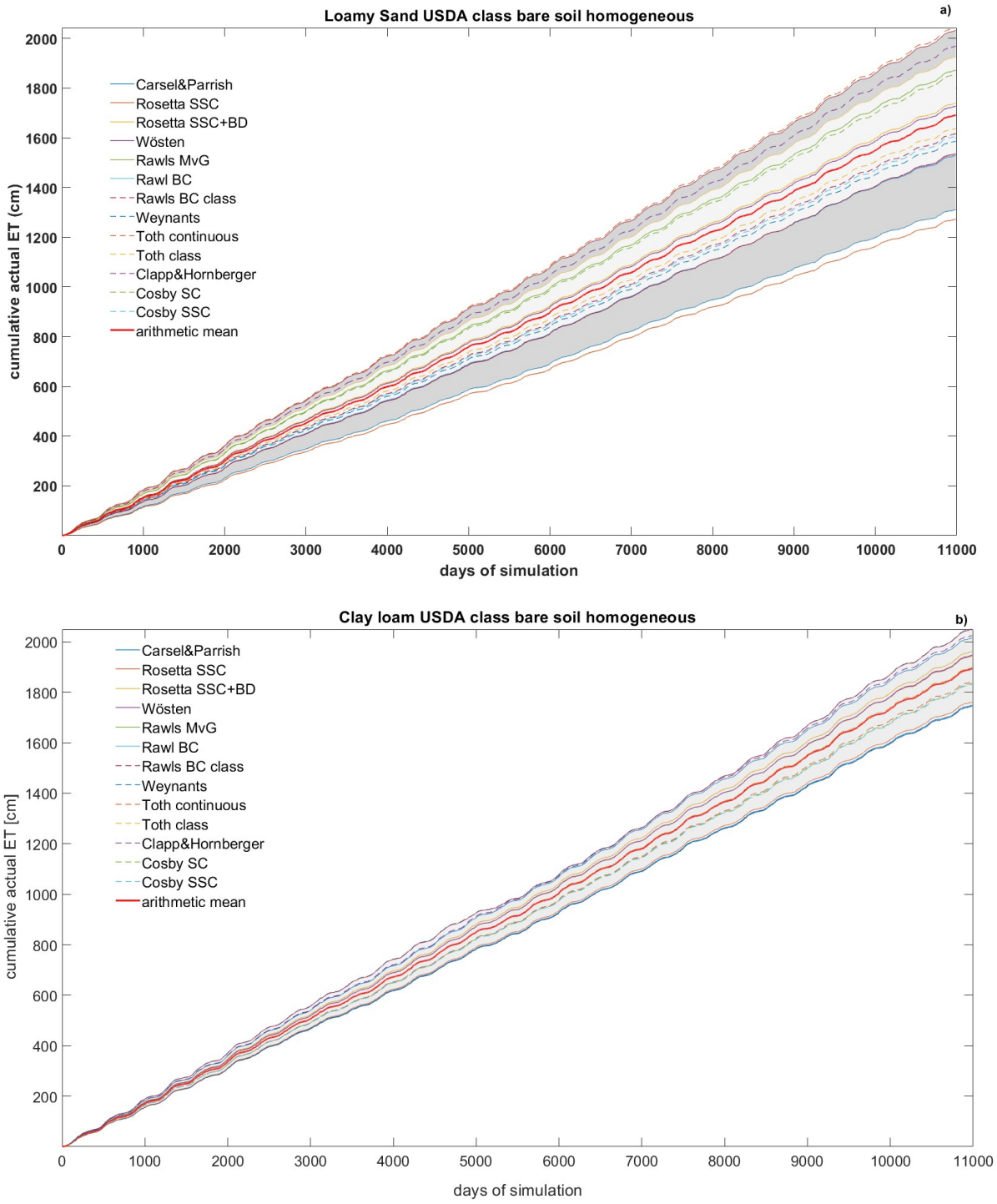


Figure 4: a) simulated cumulative actual evaporation E_a [cm] over the simulation period of 10988 days (30 years) for a bare soil with a homogeneous loamy sand soil texture. b) and a bare soil with a homogeneous clay loam soil texture. Light and dark grey shaded area represent the 70 and 90 % tolerance interval, respectively

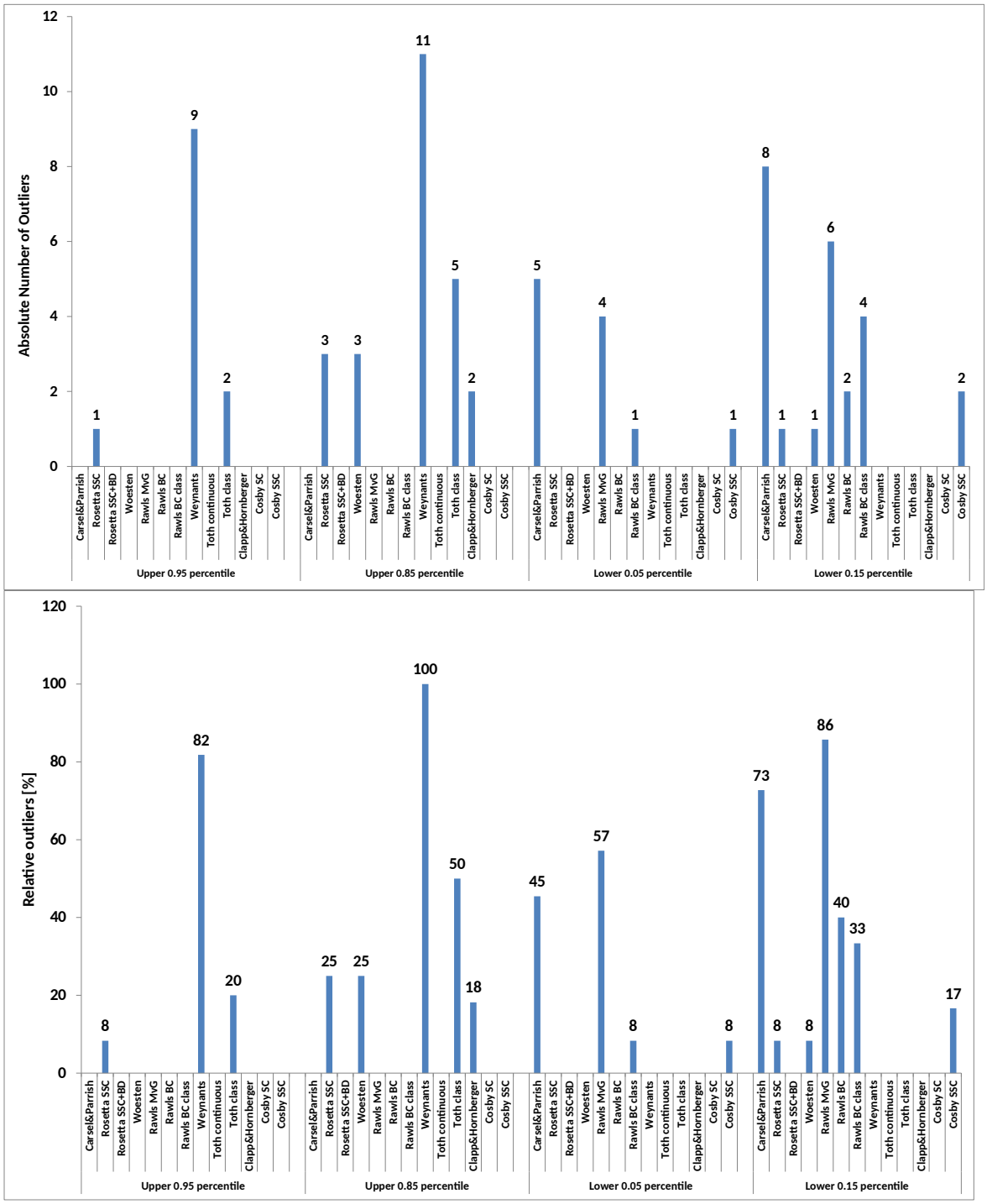


Figure 5: (a) absolute and (b) relative number of outliers for simulated E_a/ET_a at t_{end} (10988 days) for the 13 pedotransfer functions over all 12 USDA soil classes (11 for Clapp&Hornberger) and the homogeneous bare soil scenario. As some simulation runs did not converge, the total number of outliers was normalized to the number of converged simulations for each scenario and PTF combination.

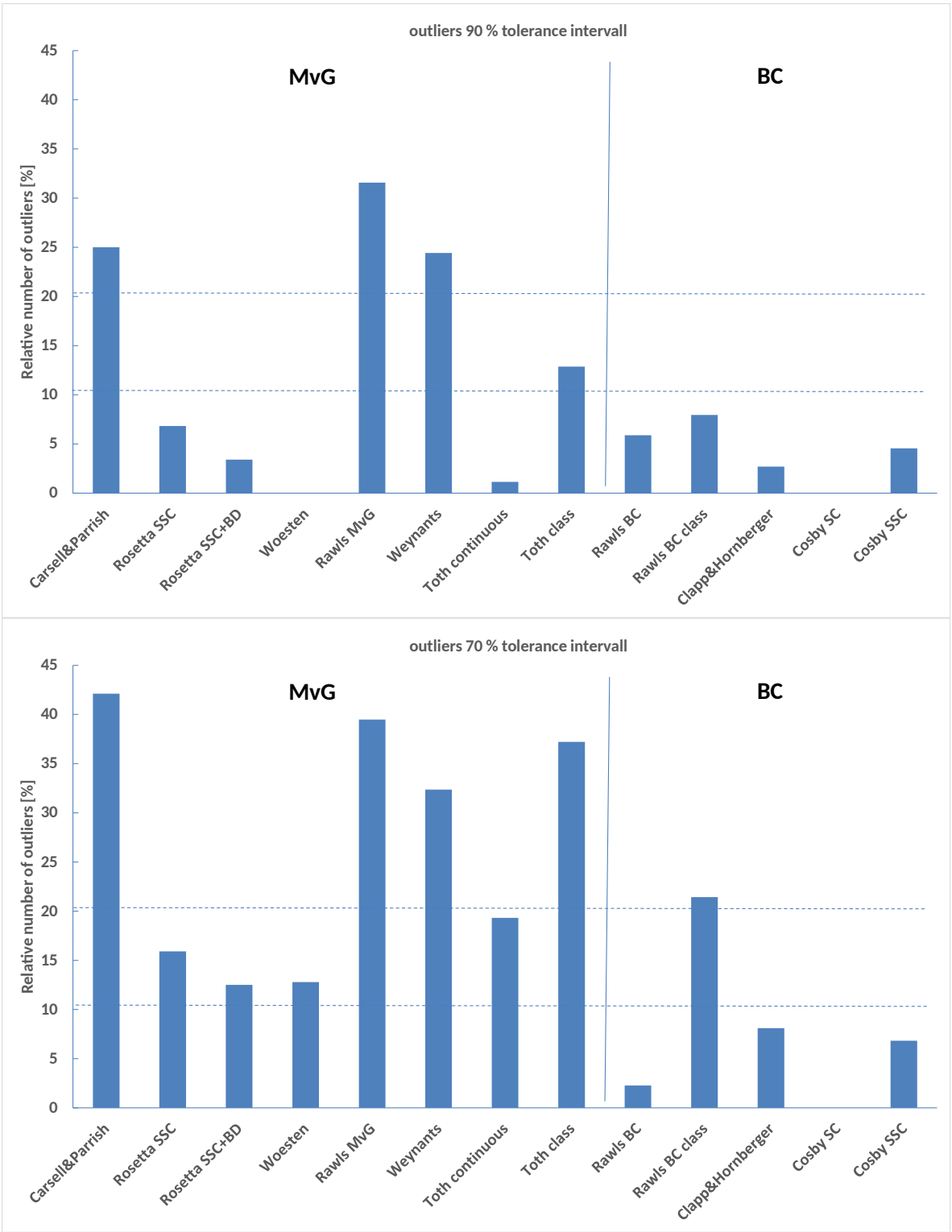
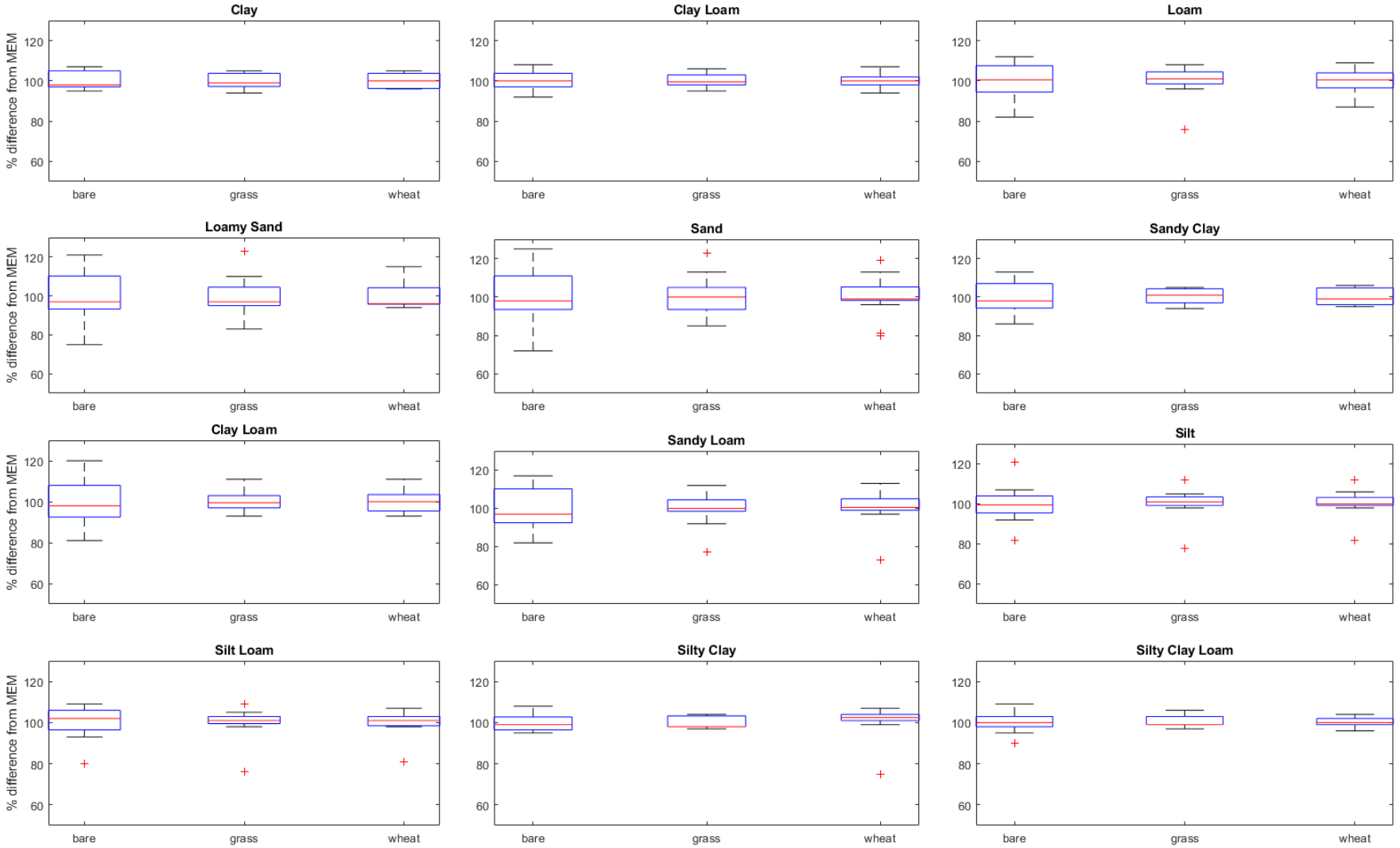


Figure 6: a) relative number of outliers for the 90 % tolerance interval and b) 70 % tolerance interval for simulated E_a or ET_a at t_{end} (10988 days) for the 13 pedotransfer functions over all USDA soil classes and scenarios depicted in Fig. 2. Dotted lines represent a threshold of 10 and 20 % outliers separating robust (<10 %), intermediate (10-20%), and non-robust (>20%) pedotransfer functions.



1420 **Figure 7:** Boxplots of relative difference in % from model ensemble mean (MEM) of simulated actual evapotranspiration, ET_a, at t_{end} = 10988 for the homogeneous soil profiles either with bare soil (E_a only) or vegetated with grass or wheat. Red line indicate the median, box the 25 and 75 percentile, whiskers the most extreme data points not considered as outliers, and crosses the outliers.

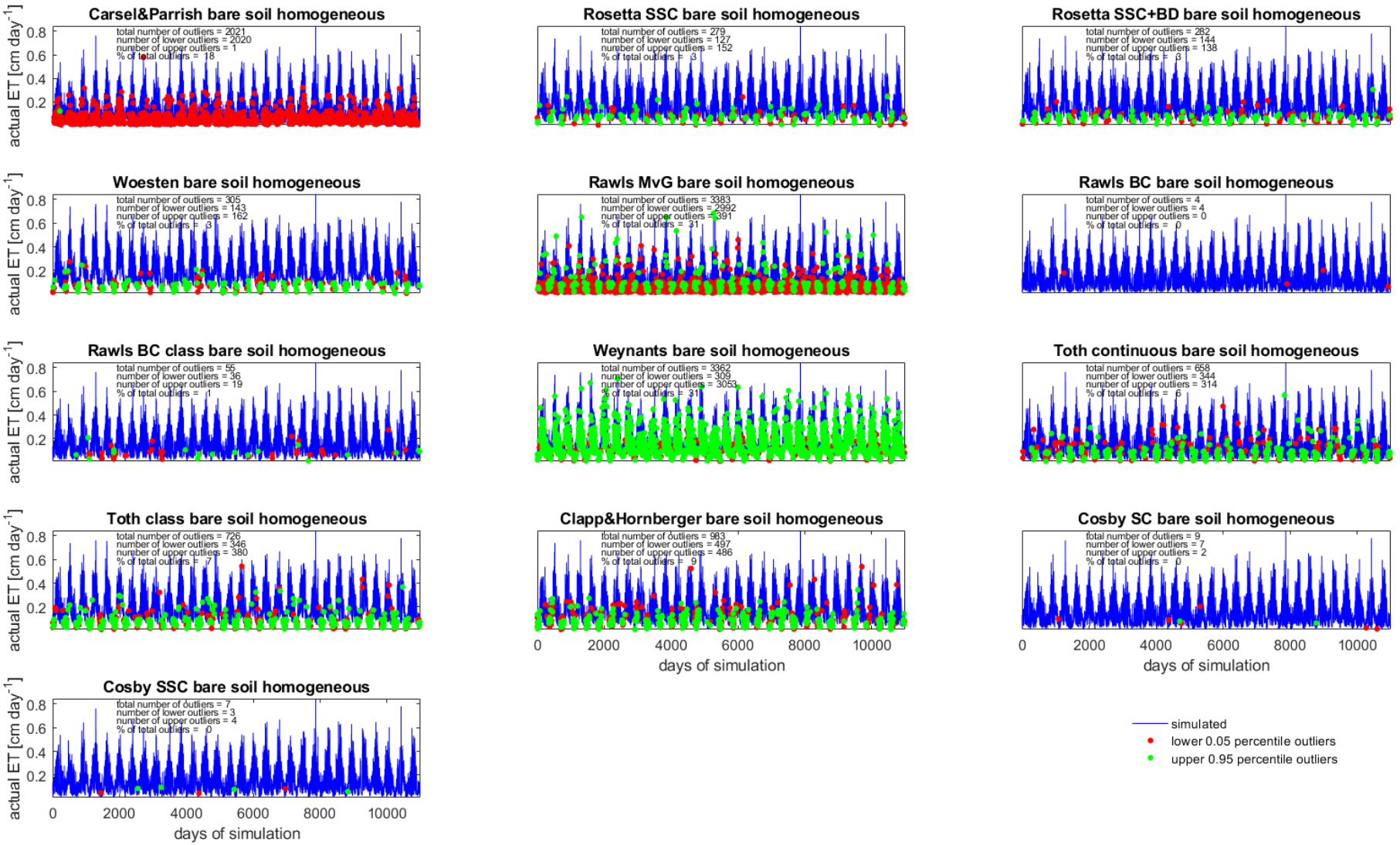


Figure 8: Actual evaporation [cm day^{-1}] for the sandy loam for the homogeneous bare soil scenario and all 13 pedotransfer functions with outliers exceeding the 90 % tolerance interval.

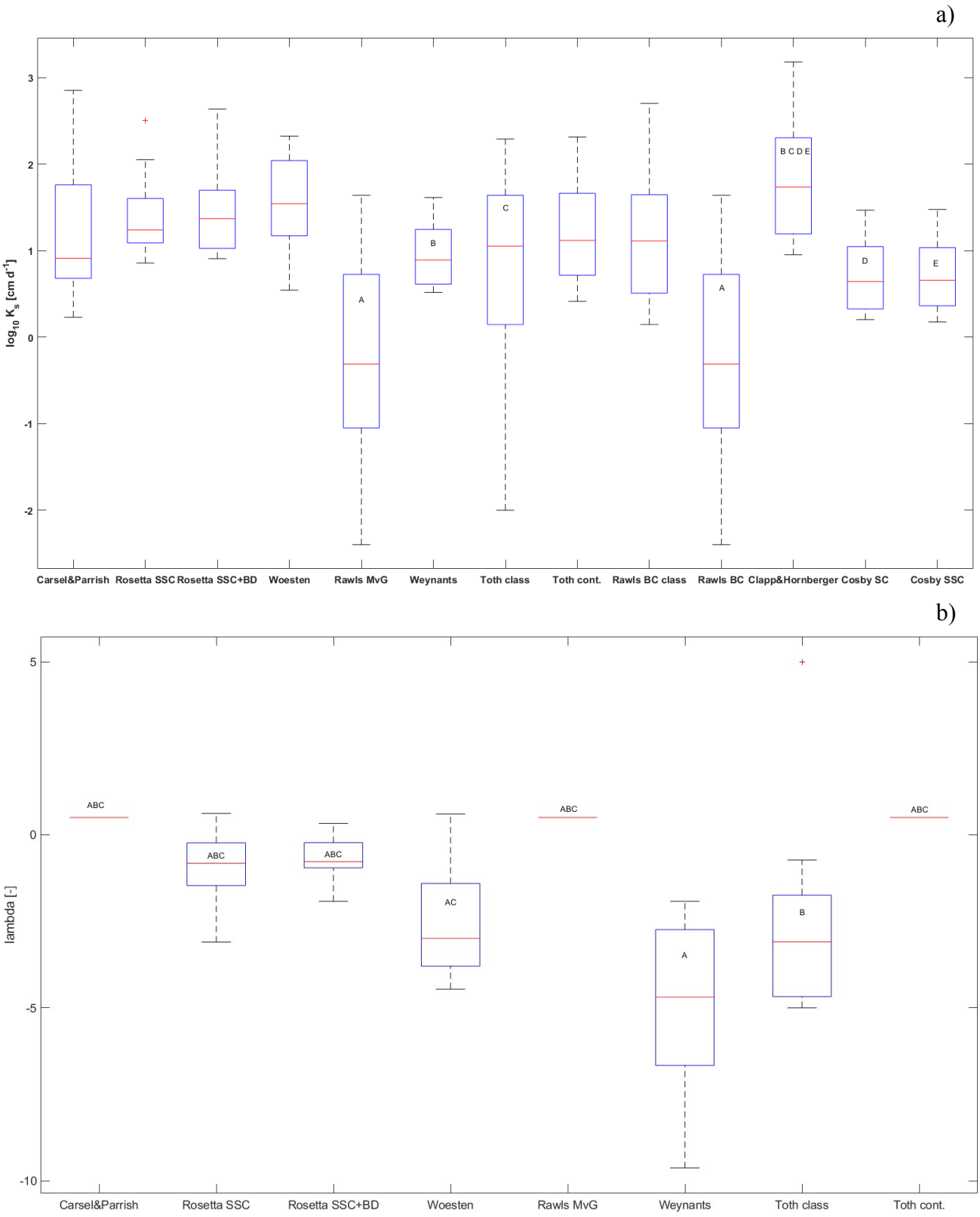


Figure 9: a) boxplots of $\log_{10} K_s$ for all PTFs, b) λ for the MvG formulation, with indication of significant differences. For significance: A differs from all other PTFs if no A is indicated, B, C, and D differ between single PTFs). Boxes are the same as for Fig. 7.

1430

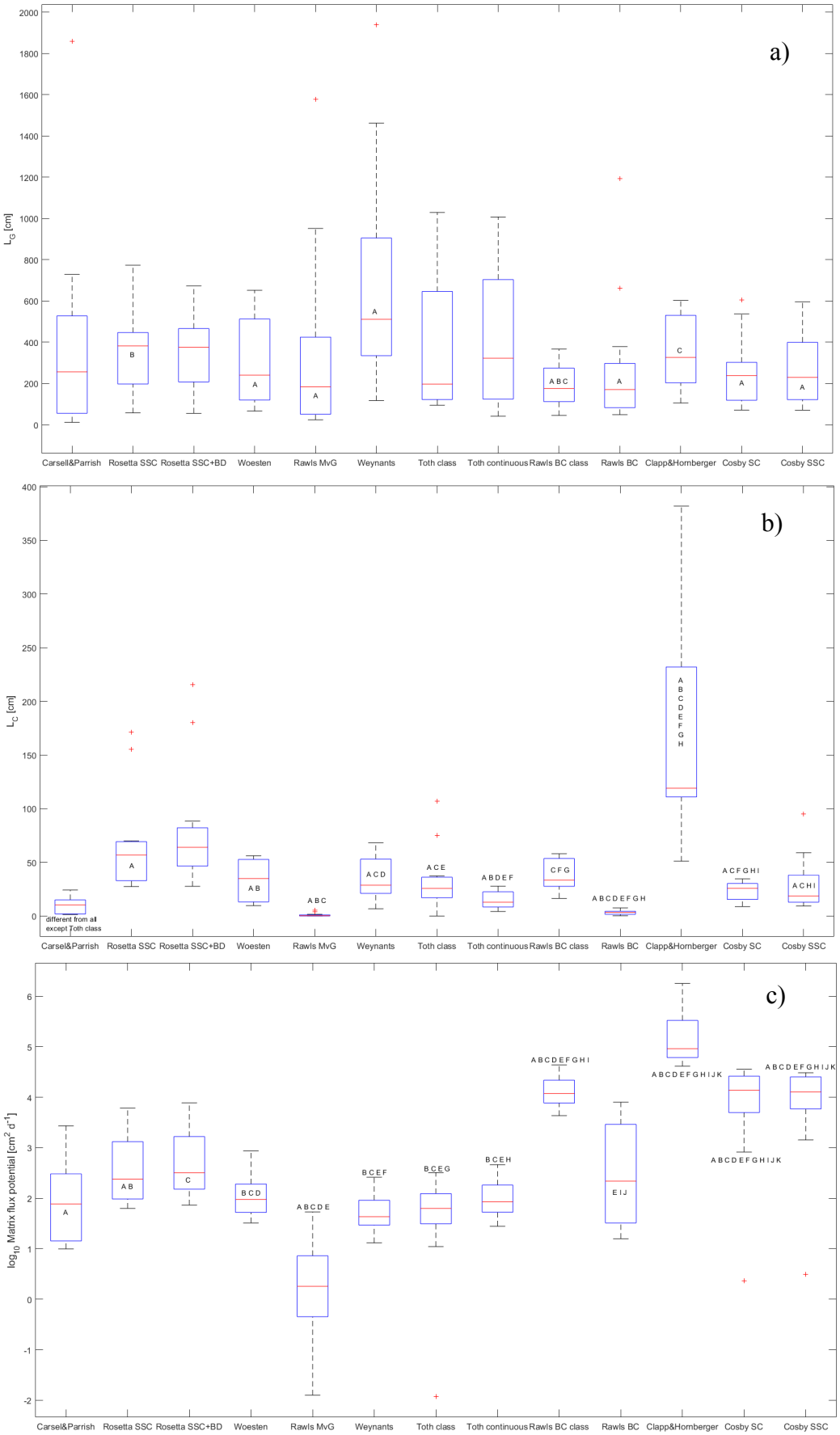


Figure 10: Boxplots for a) L_G, b) L_C, and c) $\log_{10}$ *MFP*, for all PTFs. Boxes and indication of significant differences are the same as for Fig. 7.

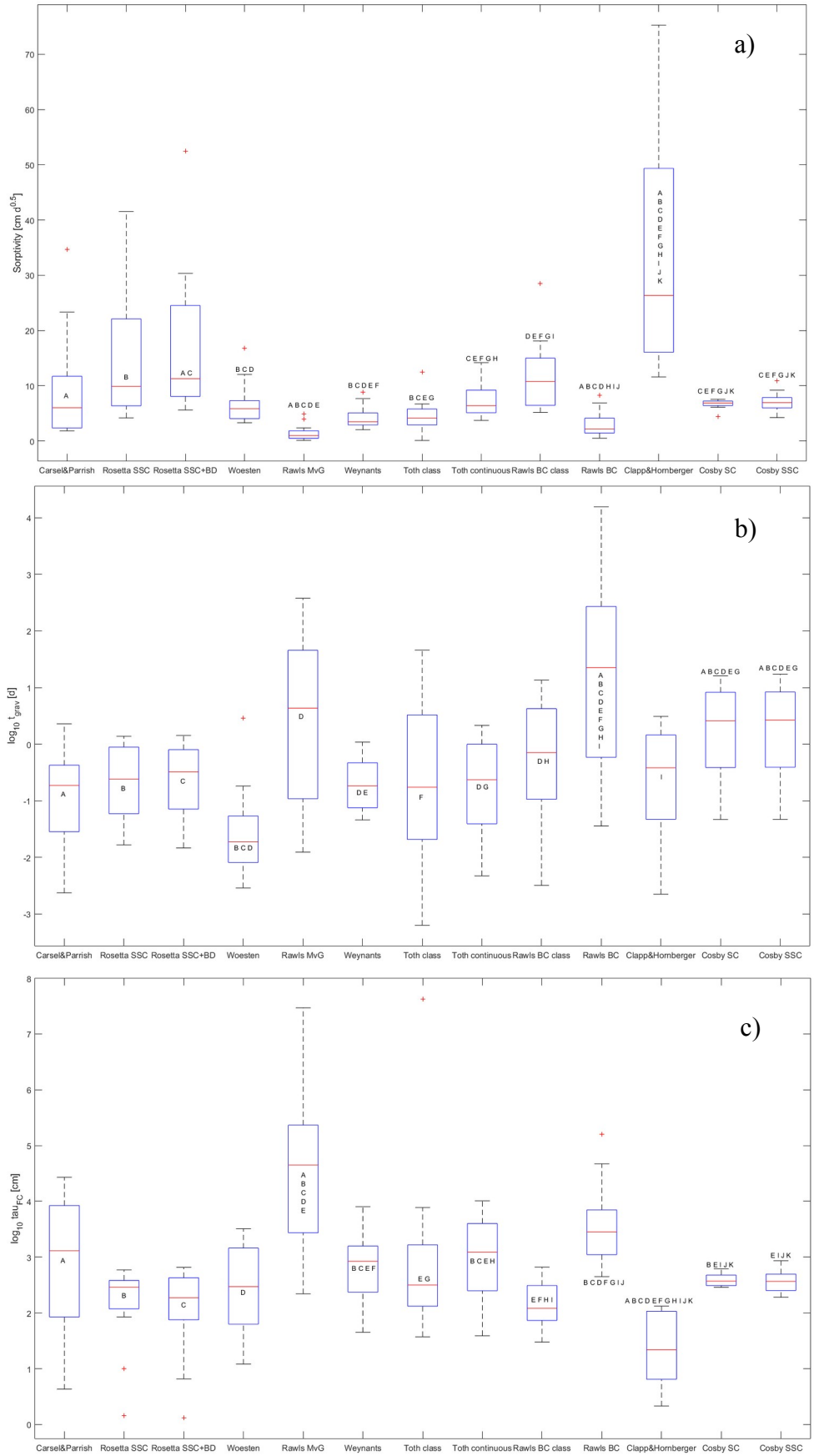
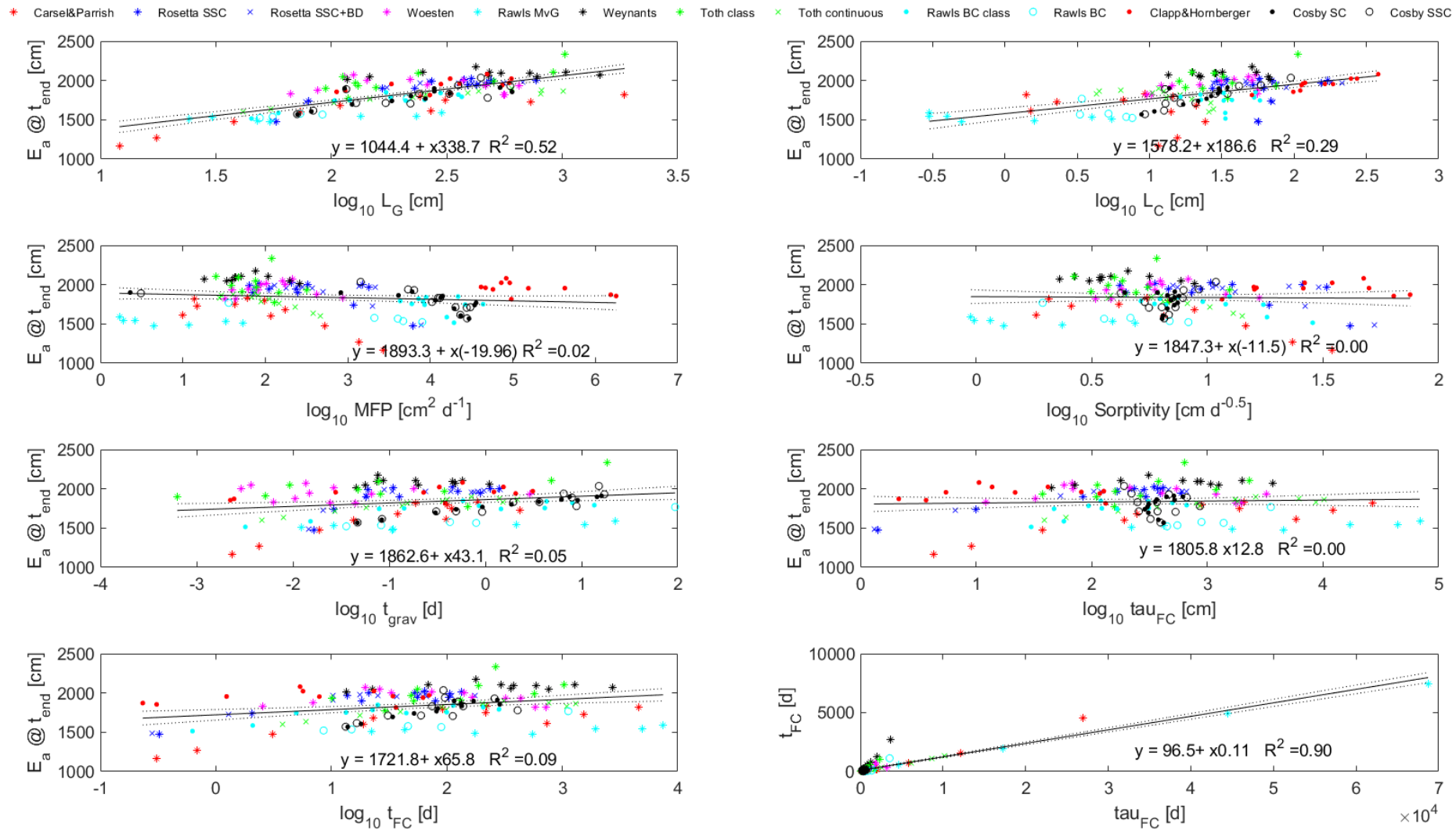


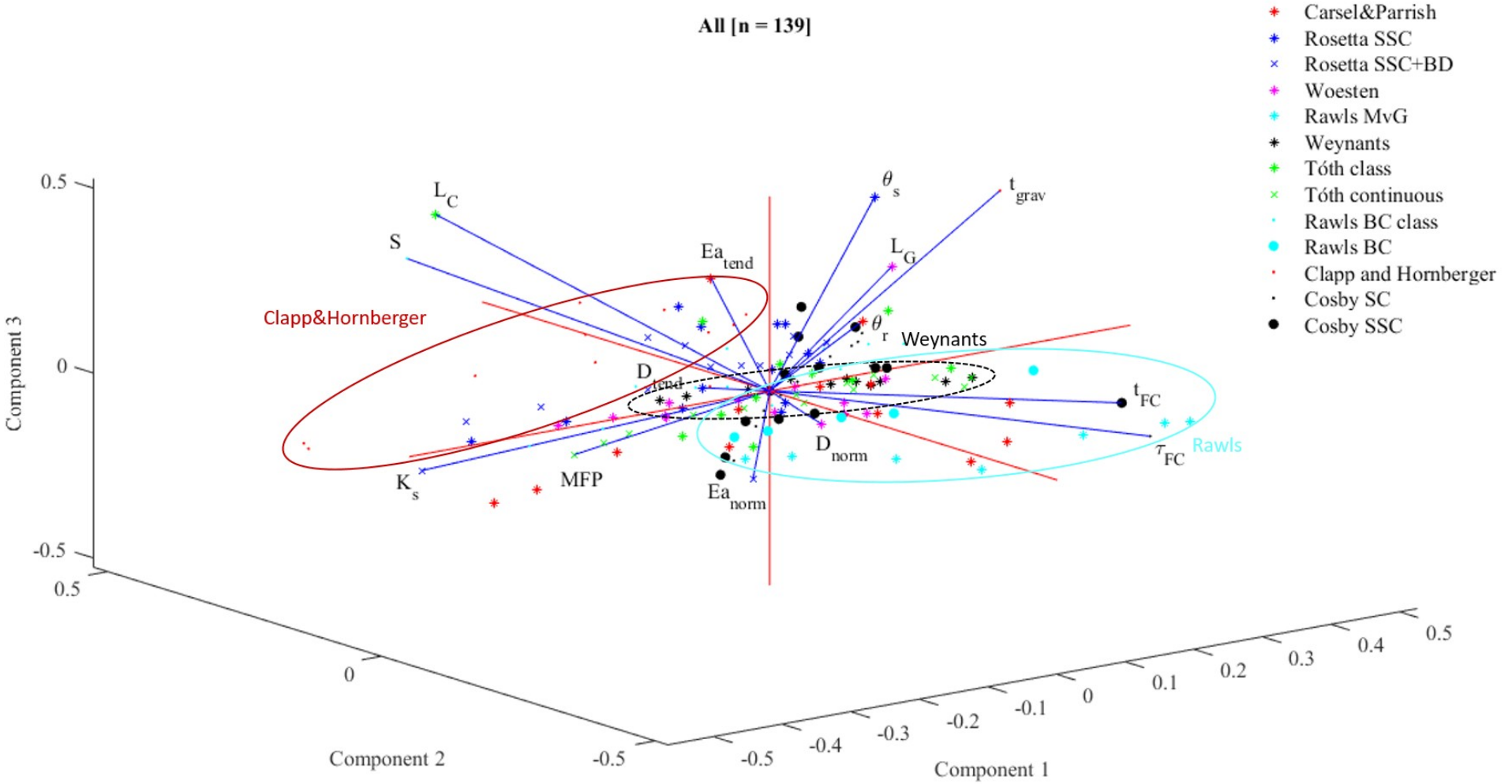
Figure 11: Boxplots for a) S b) $\log_{10} t_{grav}$ for all PTFs and c) $\log_{10} \tau_{FC}$ for all PTFs. Boxes and indication of significant differences are the same as for Fig. 7.

55



1435

Figure 12: Scatterplots of the different soil characteristics gravitational length L_G, characteristic length of evaporation L_C, matrix flux potential *MFP*, sorptivity *S*, characteristic time t_{grav}, elapsed time for the attainment of field capacity t_{FC}, characteristic time for the attainment of field capacity τ_{FC}, versus E_a at t_{end} for the homogeneous bare soil scenario as well as τ_{FC} versus t_{FC}.



1440 **Figure 13:** Trioplot of the principle component analysis for soil parameters, soil characteristics, and fluxes both available for MvG and BC. Note, that only the combination of soil parameters, soil characteristics and converged model runs were used.

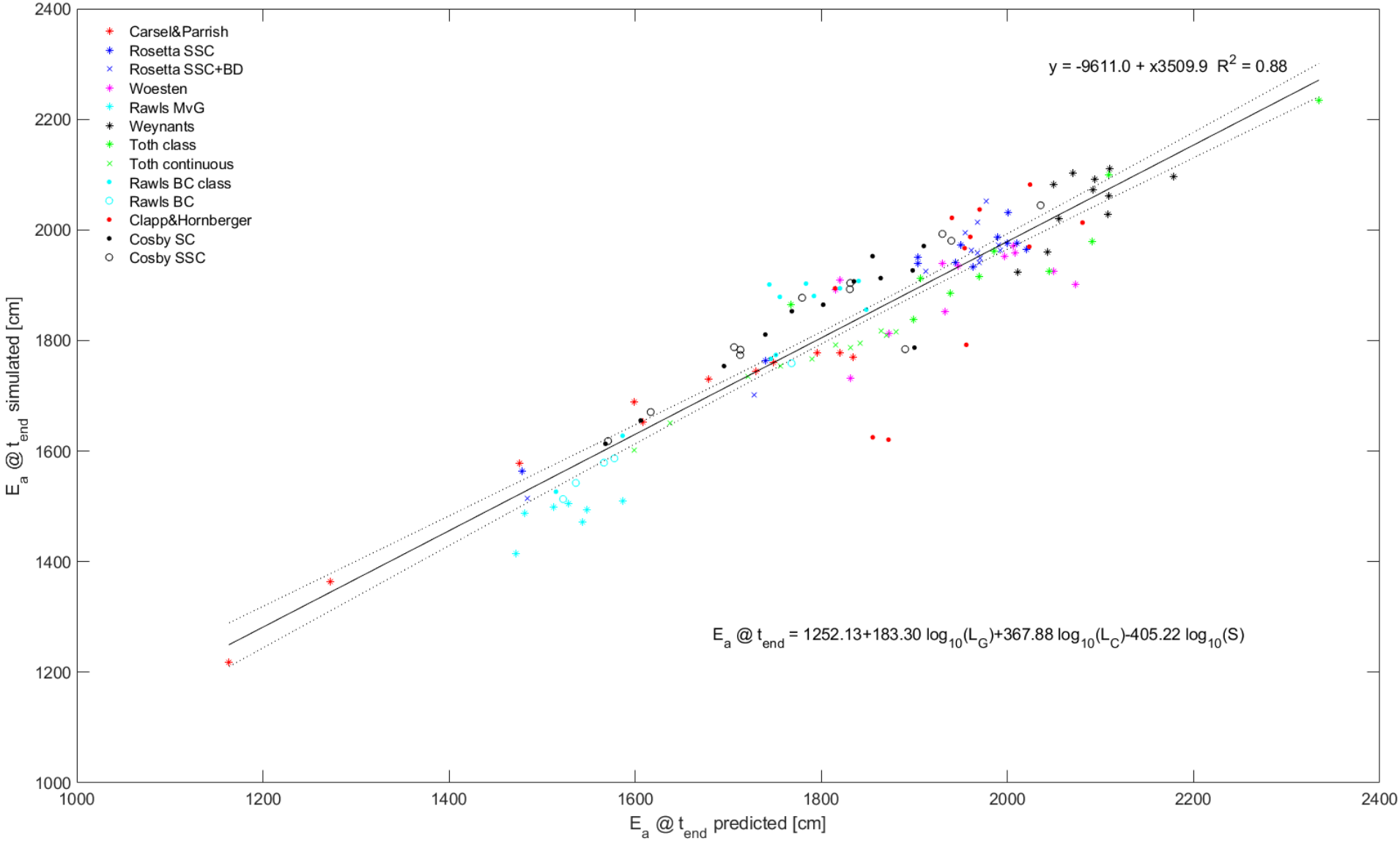


Figure 14: Predicted E_a at t_{end} [cm] by multiple regression of soil characteristics $\log_{10}(L_G)$, $\log_{10}(L_C)$, and $\log_{10}(S)$ versus simulated E_a at t_{end} [cm].

Annex

Annex Table 1: Estimated soil hydraulic parameters for the 12 USDA soil classes for the eight PTF using Mualem – van Genuchten parameterization.

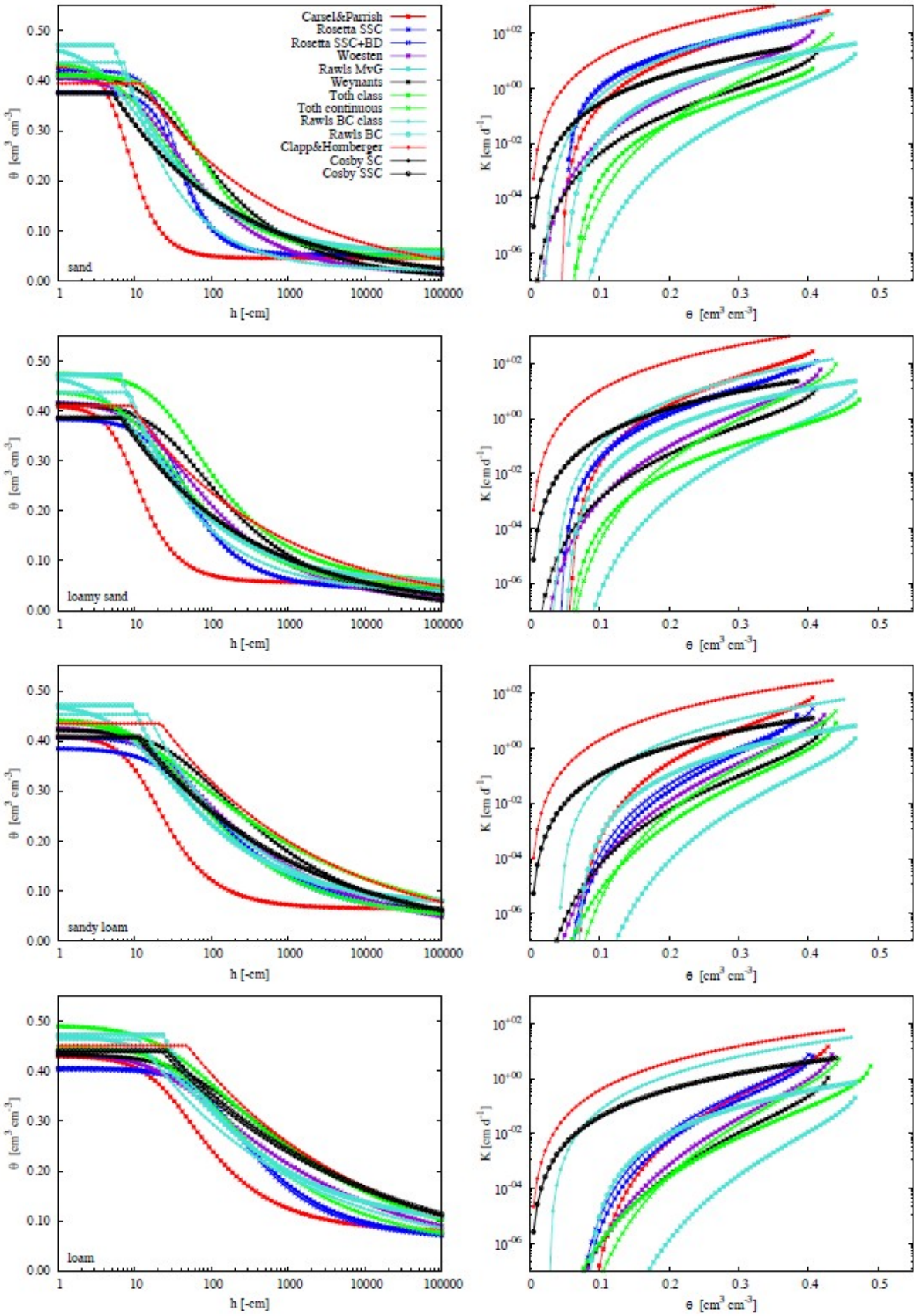
Class Clay	θ_r [$cm^3\ cm^{-3}$]	θ_s [$cm^3\ cm^{-3}$]	α [cm^{-1}]	n [-]	K_s [$cm\ d^{-1}$]	λ [-]
Carsel&Parrish (Carsel and Parrish, 1988)	0.068	0.380	0.0080	1.090	4.8	0.50
Rosetta SSC (Schaap et al., 2001)	0.097	0.485	0.0210	1.206	17.7	-2.59
Rosetta SSC+BD (Schaap et al., 2001)	0.010	0.475	0.0191	1.252	11.4	-1.83
Woesten (Wösten et al., 1999)	0.010	0.458	0.0230	1.108	24.7	-3.18
Rawls MvG (Rawls and Brakensiek, 1985)	0.101	0.472	0.0086	1.101	0.004	0.50
Weynants (Weynants et al., 2009)	0.000	0.484	0.0091	1.073	4.8	-9.63
Tóth class (Tóth et al., 2015)	0.000	0.499	0.0234	1.120	17.1	-5.00
Tóth continuous (Tóth et al. 2015)	0.041	0.456	0.0092	1.226	3.0	0.50
Class Clay loam						
Carsel&Parrish (Carsel and Parrish, 1988)	0.095	0.410	0.0190	1.310	6.2	0.50
Rosetta SSC (Schaap et al., 2001)	0.082	0.438	0.0127	1.401	7.2	-0.65
Rosetta SSC+BD (Schaap et al., 2001)	0.083	0.439	0.0118	1.438	9.9	-0.54
Woesten (Wösten et al., 1999)	0.010	0.443	0.0396	1.144	34.8	-3.87
Rawls MvG (Rawls and Brakensiek, 1985)	0.105	0.472	0.0276	1.243	0.2	0.50
Weynants (Weynants et al., 2009)	0.000	0.450	0.0139	1.136	6.9	-6.00
Tóth class (Tóth et al., 2015)	0.000	0.465	0.1284	1.116	195.2	-5.00
Tóth continuous (Tóth et al. 2015)	0.041	0.451	0.0162	1.282	9.6	0.50
Class Loam						
Carsel&Parrish (Carsel and Parrish, 1988)	0.078	0.430	0.0360	1.560	25.0	0.50
Rosetta SSC (Schaap et al., 2001)	0.063	0.406	0.0100	1.497	9.9	-0.24
Rosetta SSC+BD (Schaap et al., 2001)	0.062	0.403	0.0094	1.529	13.7	-0.21
Woesten (Wösten et al., 1999)	0.010	0.435	0.0361	1.204	35.0	-2.81
Rawls MvG (Rawls and Brakensiek, 1985)	0.082	0.472	0.0428	1.321	0.8	0.50
Weynants (Weynants et al., 2009)	0.000	0.433	0.0176	1.187	8.8	-4.23
Tóth class (Tóth et al., 2015)	0.000	0.491	0.0347	1.193	14.2	-4.30
Tóth continuous (Tóth et al. 2015)	0.041	0.449	0.0226	1.319	18.3	0.50
Class Loamy sand						
Carsel&Parrish (Carsel and Parrish, 1988)	0.057	0.410	0.1240	2.280	350.2	0.50
Rosetta SSC (Schaap et al., 2001)	0.043	0.383	0.0391	1.778	112.8	-0.89
Rosetta SSC+BD (Schaap et al., 2001)	0.046	0.415	0.0384	1.831	183.7	-0.80
Woesten (Wösten et al., 1999)	0.010	0.418	0.0579	1.390	117.8	-1.09
Rawls MvG (Rawls and Brakensiek, 1985)	0.051	0.472	0.1548	1.427	23.9	0.50
Weynants (Weynants et al., 2009)	0.000	0.415	0.0359	1.367	32.2	-2.10
Tóth class (Tóth et al., 2015)	0.052	0.475	0.0341	1.485	9.00	-1.87
Tóth continuous (Tóth et al. 2015)	0.041	0.440	0.0819	1.462	144.8	0.50
Class Sand						
Carsel&Parrish (Carsel and Parrish, 1988)	0.045	0.430	0.1450	2.680	712.8	0.50
Rosetta SSC (Schaap et al., 2001)	0.052	0.377	0.0332	2.503	322.0	-0.87
Rosetta SSC+BD (Schaap et al., 2001)	0.052	0.421	0.0347	2.495	432.6	-0.81
Woesten (Wösten et al., 1999)	0.010	0.408	0.0647	1.489	210.8	-1.02
Rawls MvG (Rawls and Brakensiek, 1985)	0.050	0.472	0.1905	1.444	43.7	0.50
Weynants (Weynants et al., 2009)	0.000	0.414	0.0404	1.424	41.2	-1.92
Tóth class (Tóth et al., 2015)	0.061	0.411	0.0258	1.801	8.3	-0.73
Tóth continuous (Tóth et al. 2015)	0.041	0.438	0.1031	1.491	206.1	0.50
Class Sandy clay						
Carsel&Parrish (Carsel and Parrish, 1988)	0.100	0.380	0.0270	1.230	2.9	0.50
Rosetta SSC (Schaap et al., 2001)	0.083	0.403	0.0290	1.206	17.0	-3.10
Rosetta SSC+BD (Schaap et al., 2001)	0.090	0.458	0.0249	1.263	23.9	-1.92
Woesten (Wösten et al., 1999)	0.010	0.432	0.0670	1.159	118.0	-4.46
Rawls MvG (Rawls and Brakensiek, 1985)	0.120	0.472	0.0750	1.172	0.8	0.50
Weynants (Weynants et al., 2009)	0.000	0.461	0.0167	1.117	12.8	-7.05
Tóth class (Tóth et al., 2015)	0.192	0.523	0.0351	1.446	43.8	-1.62
Tóth continuous (Tóth et al. 2015)	0.041	0.448	0.0261	1.310	17.5	0.50

Class Sandy clay loam	θ_r [$cm^3\ cm^{-3}$]	θ_s [$cm^3\ cm^{-3}$]	α [cm^{-1}]	n [-]	K_s [$cm\ d^{-1}$]	λ [-]
Carsel&Parrish (Carsel and Parrish, 1988)	0.100	0.390	0.0590	1.480	31.4	0.50
Rosetta SSC (Schaap et al., 2001)	0.070	0.386	0.0269	1.290	12.5	-1.79
Rosetta SSC+BD (Schaap et al., 2001)	0.075	0.441	0.0216	1.370	32.0	-1.07
Woesten (Wösten et al., 1999)	0.010	0.430	0.0746	1.185	102.9	-4.00
Rawls MvG (Rawls and Brakensiek, 1985)	0.107	0.472	0.1241	1.257	4.0	0.50
Weynants (Weynants et al., 2009)	0.000	0.443	0.0214	1.172	16.3	-5.15
Tóth class (Tóth et al., 2015)	0.000	0.409	0.0700	1.134	43.6	-5.00
Tóth continuous (Tóth et al. 2015)	0.041	0.446	0.0368	1.353	34.2	0.50
Class Sandy loam						
Carsel&Parrish (Carsel and Parrish, 1988)	0.065	0.410	0.0750	1.890	106.1	0.50
Rosetta SSC (Schaap et al., 2001)	0.045	0.385	0.0300	1.396	35.4	-1.14
Rosetta SSC+BD (Schaap et al., 2001)	0.047	0.408	0.0258	1.460	54.7	-0.84
Woesten (Wösten et al., 1999)	0.010	0.427	0.0510	1.278	70.8	-2.00
Rawls MvG (Rawls and Brakensiek, 1985)	0.067	0.472	0.1087	1.372	7.1	0.50
Weynants (Weynants et al., 2009)	0.000	0.423	0.0268	1.261	19.0	-3.01
Tóth class (Tóth et al., 2015)	0.000	0.441	0.0750	1.190	44.9	-4.35
Tóth continuous (Tóth et al. 2015)	0.041	0.443	0.0485	1.397	62.2	0.50
Class Silt						
Carsel&Parrish (Carsel and Parrish, 1988)	0.034	0.460	0.0160	1.370	6.0	0.50
Rosetta SSC (Schaap et al., 2001)	0.049	0.510	0.0074	1.664	45.6	0.62
Rosetta SSC+BD (Schaap et al., 2001)	0.053	0.427	0.0063	1.657	45.7	0.33
Woesten (Wösten et al., 1999)	0.010	0.423	0.0075	1.406	3.5	0.60
Rawls MvG (Rawls and Brakensiek, 1985)	0.020	0.472	0.0178	1.357	0.2	0.50
Weynants (Weynants et al., 2009)	0.000	0.414	0.0131	1.283	3.3	-2.47
Tóth class (Tóth et al., 2015)	0.000	0.465	0.0042	1.485	1.4	-2.64
Tóth continuous (Tóth et al. 2015)	0.041	0.453	0.0106	1.275	7.1	0.50
Class Silt loam						
Carsel&Parrish (Carsel and Parrish, 1988)	0.067	0.450	0.0200	1.410	10.8	0.50
Rosetta SSC (Schaap et al., 2001)	0.061	0.430	0.0044	1.684	21.9	0.46
Rosetta SSC+BD (Schaap et al., 2001)	0.059	0.398	0.0051	1.675	23.0	0.22
Woesten (Wösten et al., 1999)	0.010	0.431	0.0183	1.252	13.2	-1.72
Rawls MvG (Rawls and Brakensiek, 1985)	0.062	0.472	0.0236	1.340	0.3	0.50
Weynants (Weynants et al., 2009)	0.000	0.426	0.0144	1.210	5.1	-3.96
Tóth class (Tóth et al., 2015)	0.000	0.424	0.0074	1.255	1.2	-3.55
Tóth continuous (Tóth et al. 2015)	0.041	0.452	0.0143	1.288	9.9	0.50
Class Silty clay						
Carsel&Parrish (Carsel and Parrish, 1988)	0.068	0.380	0.0080	1.090	4.8	0.50
Rosetta SSC (Schaap et al., 2001)	0.101	0.500	0.0130	1.364	15.3	-0.77
Rosetta SSC+BD (Schaap et al., 2001)	0.097	0.475	0.0127	1.372	8.1	-0.75
Woesten (Wösten et al., 1999)	0.010	0.453	0.0217	1.115	15.6	-3.72
Rawls MvG (Rawls and Brakensiek, 1985)	0.105	0.472	0.0096	1.169	0.01	0.50
Weynants (Weynants et al., 2009)	0.000	0.466	0.0088	1.108	3.3	-7.87
Tóth class (Tóth et al., 2015)	0.000	0.455	0.0309	1.111	0.01	5.00
Tóth continuous (Tóth et al. 2015)	0.041	0.457	0.0074	1.223	2.6	0.50
Class Silty clay loam						
Carsel&Parrish (Carsel and Parrish, 1988)	0.089	0.430	0.0100	1.230	1.7	0.50
Rosetta SSC (Schaap et al., 2001)	0.091	0.477	0.0087	1.492	12.1	-0.22
Rosetta SSC+BD (Schaap et al., 2001)	0.088	0.456	0.0085	1.498	9.1	-0.24
Woesten (Wösten et al., 1999)	0.010	0.445	0.0221	1.144	14.2	-3.35
Rawls MvG (Rawls and Brakensiek, 1985)	0.099	0.472	0.0130	1.237	0.04	0.50
Weynants (Weynants et al., 2009)	0.000	0.451	0.0101	1.141	3.5	-6.28
Tóth class (Tóth et al., 2015)	0.000	0.463	0.0107	1.189	1.4	-2.64
Tóth continuous (Tóth et al. 2015)	0.041	0.456	0.0087	1.240	3.8	0.50

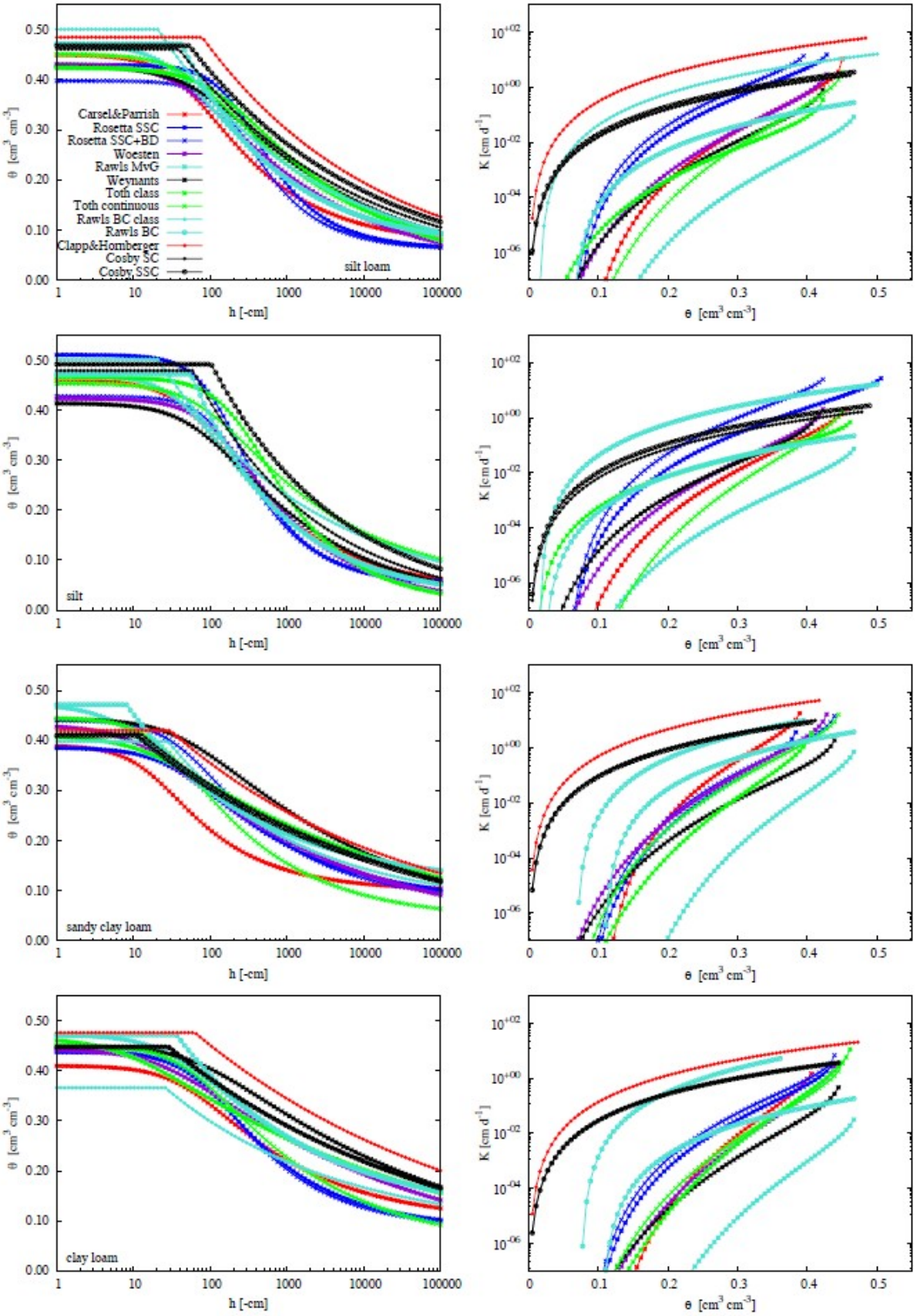
Annex Table 2: Estimated soil hydraulic parameters for the 12 USDA soil classes for the five PTF using Brooks and Corey (Campbell in case of θ_r = 0) parameterization. Note that for the class Silt no parameters are reported (*NR*) for Clapp&Hornberger (1978). Note that n in Eq. [1] is $1/b$.

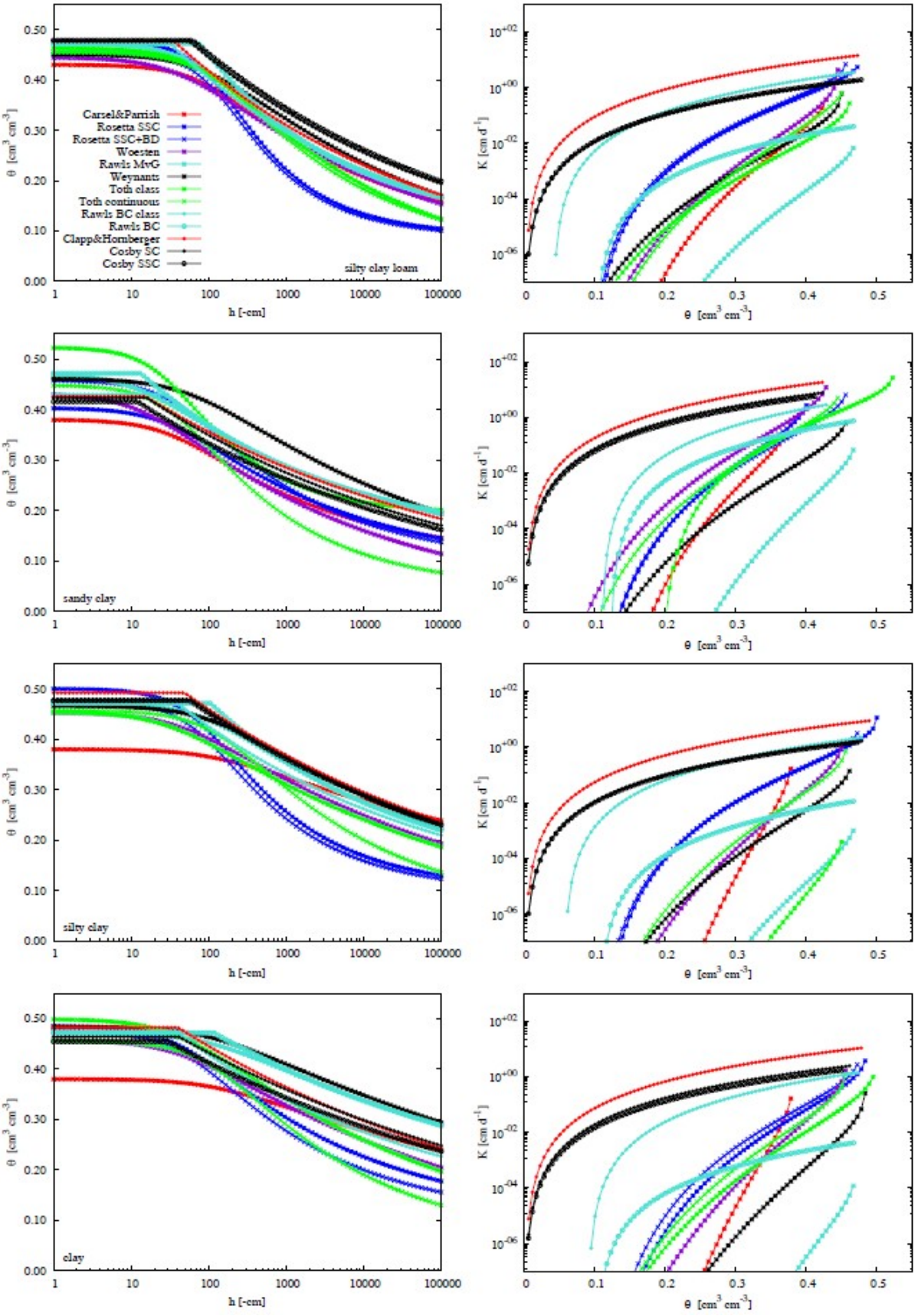
Class Clay	θ_r [cm^3 cm^{-3}]	θ_s [cm^3 cm^{-3}]	α [cm^{-1}]	b [-]	K_s [cm d^{-1}]
Rawls BC class (Rawls et al., 1982)	0.090	0.475	0.0268	7,634	1.4
Rawls BC (Rawls and Brakensiek, 1985)	0.101	0.472	0.0086	9.883	0.004
Clapp&Hornberger (Clapp and Hornberger, 1978)	0,000	0.482	0.0247	11.400	11.1
Cosby SC (Cosby et al., 1984)	0,000	0.464	0.0241	12.450	2.5
Cosby SSC (Cosby et al., 1984)	0,000	0.454	0.0334	12.460	1.8
Class Clay loam					
Rawls BC class (Rawls et al., 1982)	0.075	0.366	0.0386	5.155	5.5
Rawls BC (Rawls and Brakensiek, 1985)	0.105	0.472	0.0276	4.110	0.2
Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.476	0.0159	8.520	21.2
Cosby SC (Cosby et al., 1984)	0.000	0.449	0.0346	8.157	3.8
Cosby SSC (Cosby et al., 1984)	0.000	0.447	0.0350	8.185	3.7
Class Loam					
Rawls BC class (Rawls et al., 1982)	0.027	0.463	0.0897	4.545	31.7
Rawls BC (Rawls and Brakensiek, 1985)	0.082	0.472	0.0428	3.116	0.8
Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.451	0.0209	5.390	60.1
Cosby SC (Cosby et al., 1984)	0.000	0.439	0.0441	6.090	5.1
Cosby SSC (Cosby et al., 1984)	0.000	0.441	0.0387	6.120	5.6
Class Loamy sand					
Rawls BC class (Rawls et al., 1982)	0.035	0.437	0.1150	2.110	146.6
Rawls BC (Rawls and Brakensiek, 1985)	0.051	0.472	0.1548	2.343	23.9
Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.410	0.1111	4.380	1350.7
Cosby SC (Cosby et al., 1984)	0.000	0.386	0.1564	3.864	22.2
Cosby SSC (Cosby et al., 1984)	0.000	0.386	0.1457	3.796	23.5
Class Sand					
Rawls BC class (Rawls et al., 1982)	0.020	0.437	0.1380	1.689	504.0
Rawls BC (Rawls and Brakensiek, 1985)	0.050	0.472	0.1905	2.254	43.7
Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.395	0.0826	4.050	1520.6
Cosby SC (Cosby et al., 1984)	0.000	0.376	0.1991	3.705	29.4
Cosby SSC (Cosby et al., 1984)	0.000	0.376	0.1921	3.615	30.0
Class Sandy clay					
Rawls BC class (Rawls et al., 1982)	0.109	0.430	0.0343	5.952	2.9
Rawls BC (Rawls and Brakensiek, 1985)	0.120	0.472	0.0750	5.814	0.8
Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.426	0.0654	10.400	18.7
Cosby SC (Cosby et al., 1984)	0.000	0.423	0.0633	9.588	7.7
Cosby SSC (Cosby et al., 1984)	0.000	0.416	0.0825	9.538	5.8

Class Sandy clay loam	θ_r [cm^3 cm^{-3}]	θ_s [cm^3 cm^{-3}]	α [cm^{-1}]	b [-]	K_s [cm d^{-1}]
Rawls BC class (Rawls et al., 1982)	0.068	0.398	0.0356	4.000	10.3
Rawls BC (Rawls and Brakensiek, 1985)	0.107	0.472	0.1241	3.893	4.0
Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.420	0.0334	7.120	54.4
Cosby SC (Cosby et al., 1984)	0.000	0.413	0.0805	7.362	10.2
Cosby SSC (Cosby et al., 1984)	0.000	0.409	0.0900	7.316	9.0
Class Sandy loam					
Rawls BC class (Rawls et al., 1982)	0.041	0.453	0.0682	3.106	62.2
Rawls BC (Rawls and Brakensiek, 1985)	0.067	0.472	0.1087	2.689	7.1
Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.435	0.0459	4.900	299.5
Cosby SC (Cosby et al., 1984)	0.000	0.407	0.0936	4.818	12.2
Cosby SSC (Cosby et al., 1984)	0.000	0.408	0.0856	4.789	13.1
Class Silt					
Rawls BC class (Rawls et al., 1982)	0.015	0.501	0.0482	4.739	16.3
Rawls BC (Rawls and Brakensiek, 1985)	0.020	0.472	0.0178	2.797	0.2
Clapp&Hornberger (Clapp and Hornberger, 1978)	NR	NR	NR	NR	NR
Cosby SC (Cosby et al., 1984)	0.000	0.479	0.0168	3.705	1.6
Cosby SSC (Cosby et al., 1984)	0.000	0.492	0.0097	3.861	2.8
Class Silt loam					
Rawls BC class (Rawls et al., 1982)	0.015	0.501	0.0482	4.739	16.3
Rawls BC (Rawls and Brakensiek, 1985)	0.062	0.472	0.0236	2.939	0.3
Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.485	0.0127	5.300	62.2
Cosby SC (Cosby et al., 1984)	0.000	0.461	0.0256	5.295	2.7
Cosby SSC (Cosby et al., 1984)	0.000	0.468	0.0187	5.389	3.6
Class Silty clay					
Rawls BC class (Rawls et al., 1982)	0.056	0.479	0.0292	7.874	2.2
Rawls BC (Rawls and Brakensiek, 1985)	0.105	0.472	0.0096	5.914	0.01
Clapp&Hornberger (Clapp and Hornberger, 1978)	0,000	0.492	0.0204	10.400	9.0
Cosby SC (Cosby et al., 1984)	0,000	0.479	0.0168	1.224	1.6
Cosby SSC (Cosby et al., 1984)	0,000	0.477	0.0176	1.298	1.5
Class Silty clay loam					
Rawls BC class (Rawls et al., 1982)	0.040	0.471	0.0307	6.623	3.6
Rawls BC (Rawls and Brakensiek, 1985)	0.099	0.472	0.0130	4,226	0.04
Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.477	0.0281	7.750	14.7
Cosby SC (Cosby et al., 1984)	0.000	0.476	0.0178	8.316	1.8
Cosby SSC (Cosby et al., 1984)	0.000	0.478	0.0159	8.408	1.9

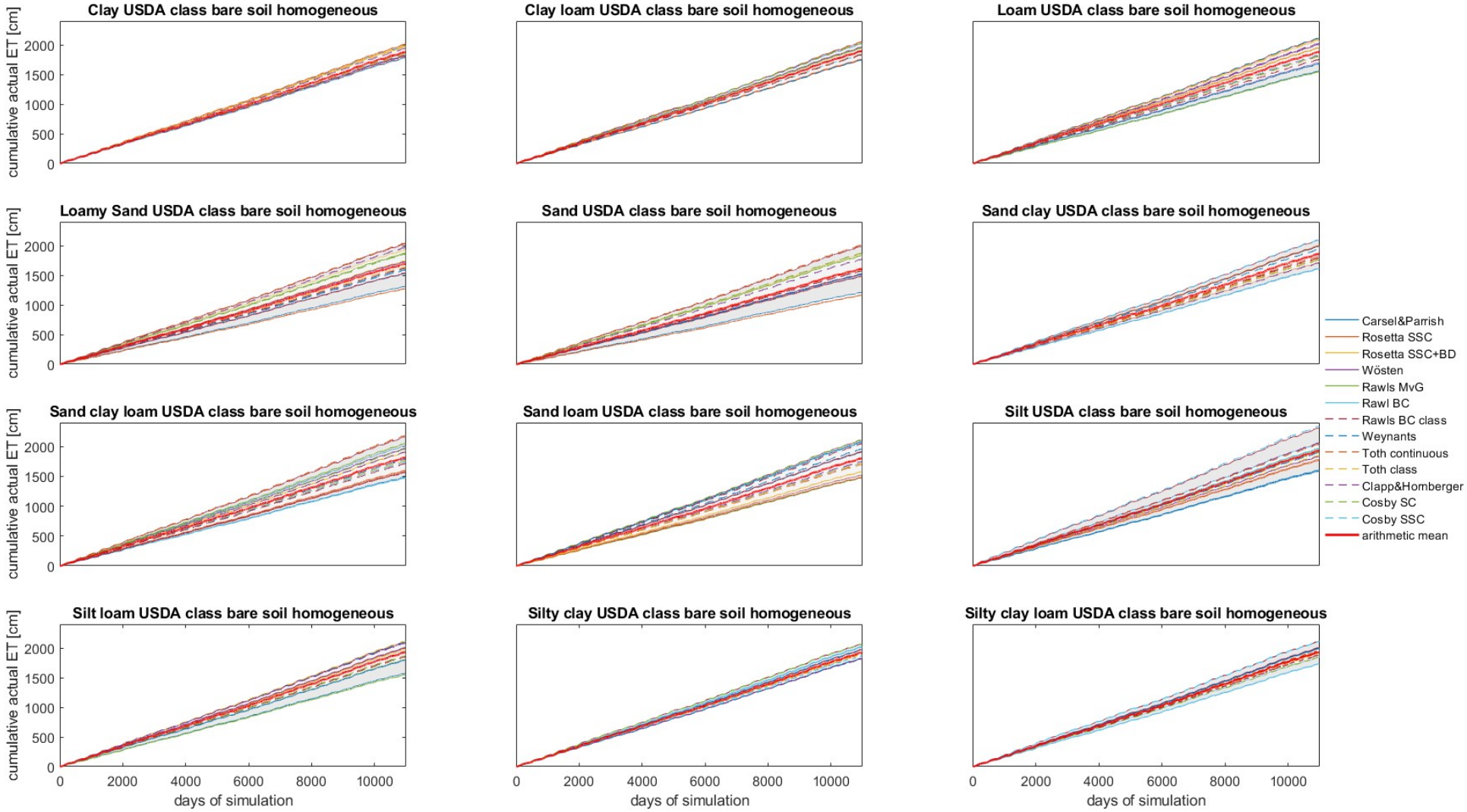

Carsel&Parrish
Rosetta SSC
Rosetta SSC+BD
Woesten
Rawls MvG
Weynants
Toth class
Toth continuous
Rawls BC class
Rawls BC
Clapp&Homberger
Cosby SC
Cosby SSC
sand
loamy sand
sandy loam
loam
θ [cm³ cm⁻³]
h [-cm]
K [cm d⁻¹]

1455


Carsel&Parrish
Rosetta SSC
Rosetta SSC+BD
Woesten
Rawls MvG
Weynants
Toth class
Toth continuous
Rawls BC class
Rawls BC
Clapp&Hornberger
Cosby SC
Cosby SSC
silt loam
silt
sandy clay loam
clay loam
θ [cm³ cm⁻³]
h [-cm]
K [cm d⁻¹]

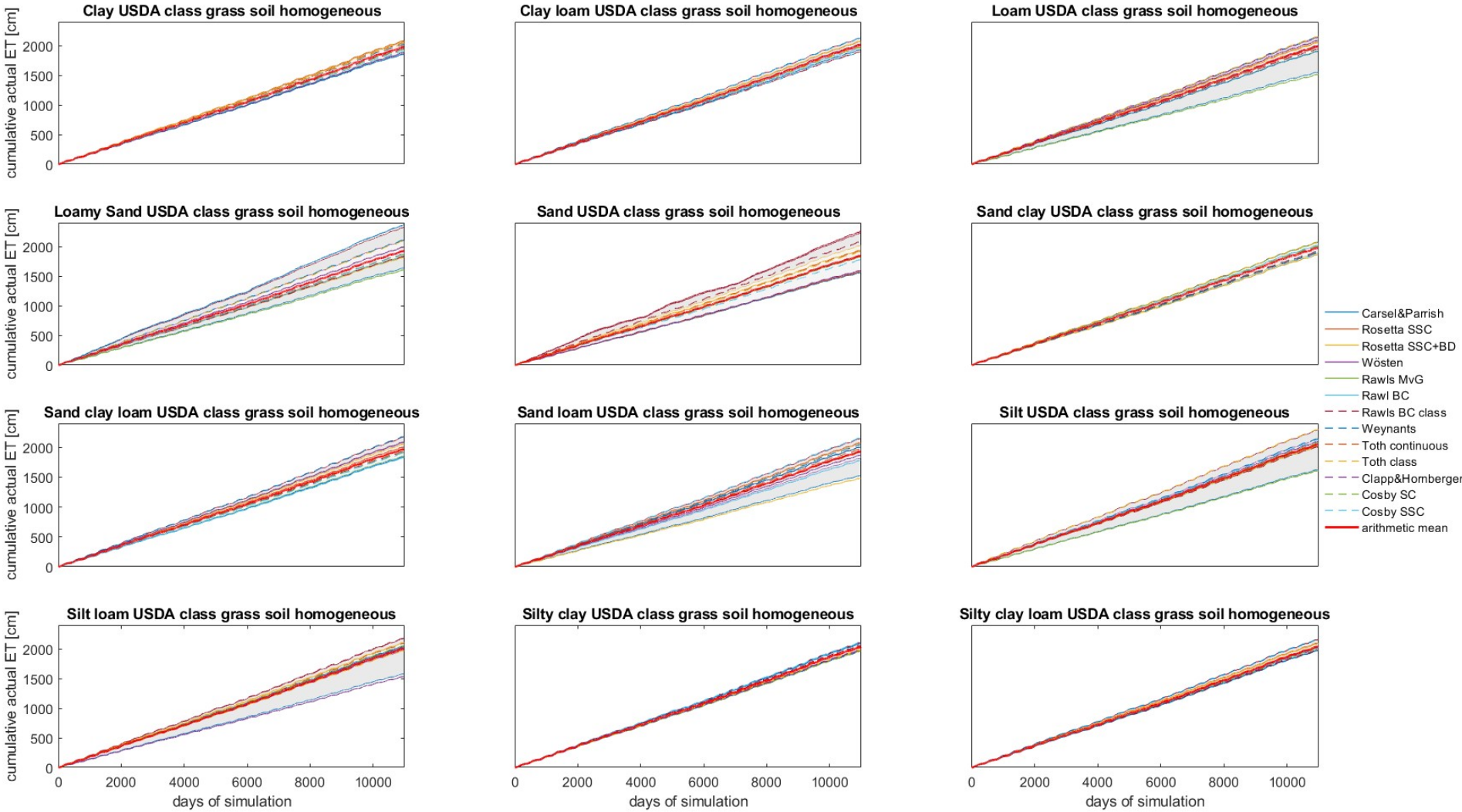


Annex Figure 1: Retention (left) and hydraulic conductivity curves right) for the for the 12 USDA soil classes for the 13 PTF (Parameters listed in Annex Tab. 1 and 2). Note that y-axis for the hydraulic conductivity is in log-scale.



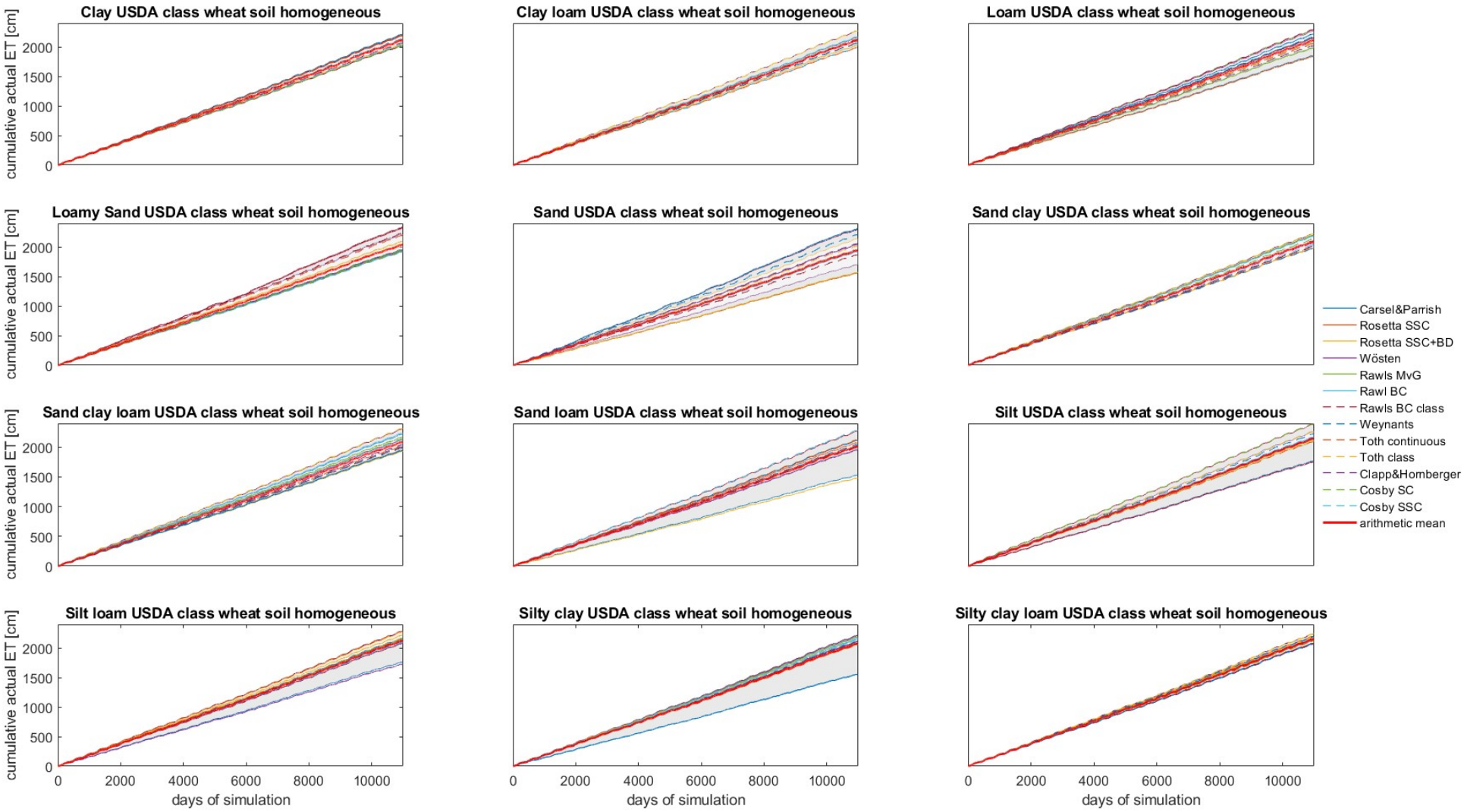
1465 **Annex Figure 2:** Simulated cumulative actual evaporation E_a [cm] over the simulation period of 10988 days (30 years) for the homogeneous bare soil scenario. Light and dark grey shaded area represent the 70 and 90 % tolerance interval, respectively.

65

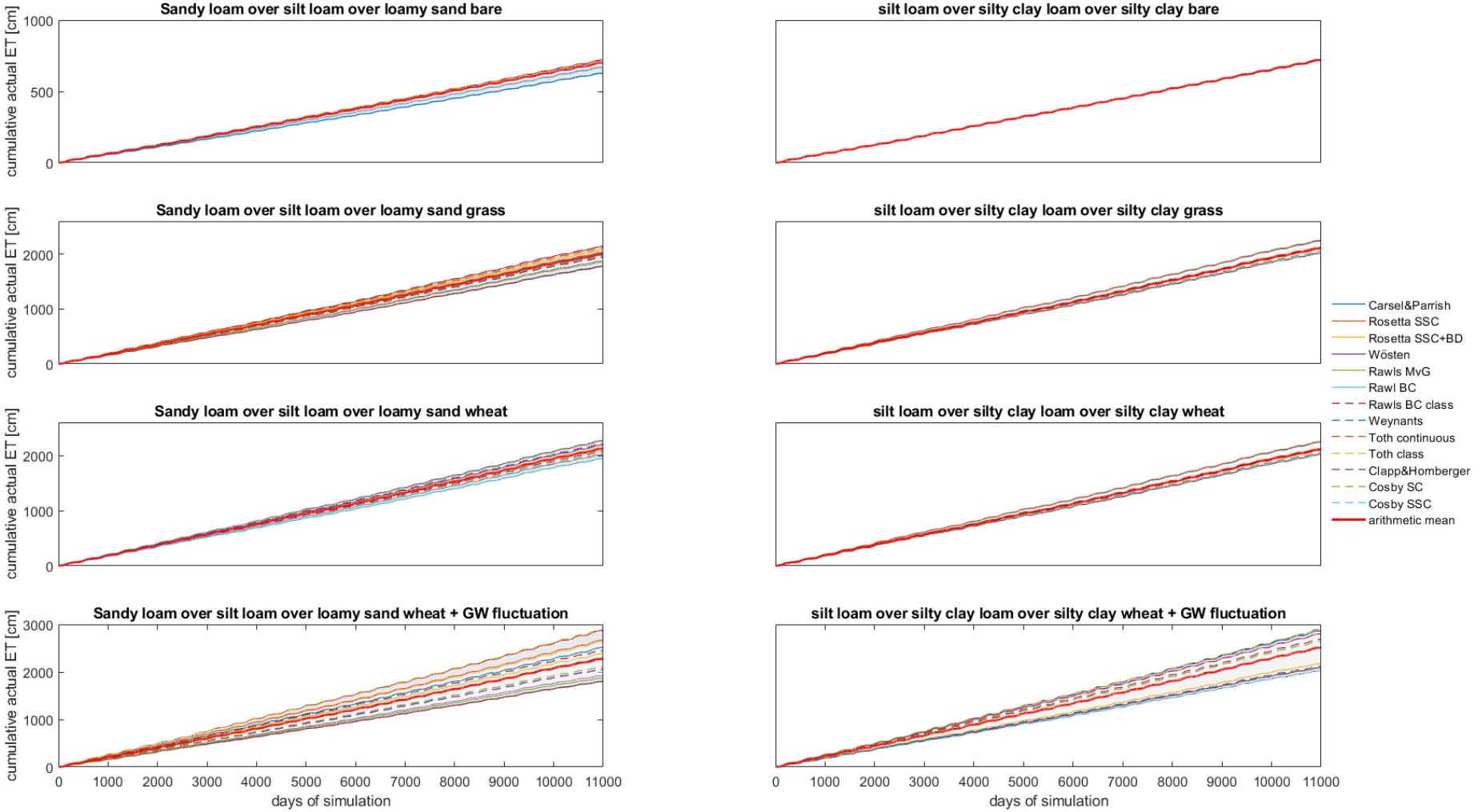


Annex Figure 3: Simulated cumulative actual evapotranspiration ET_a [cm] over the simulation period of 10988 days (30 years) for the homogeneous grass vegetated scenario. Light and dark grey shaded area represent the 70 and 90 % tolerance interval, respectively.

1470

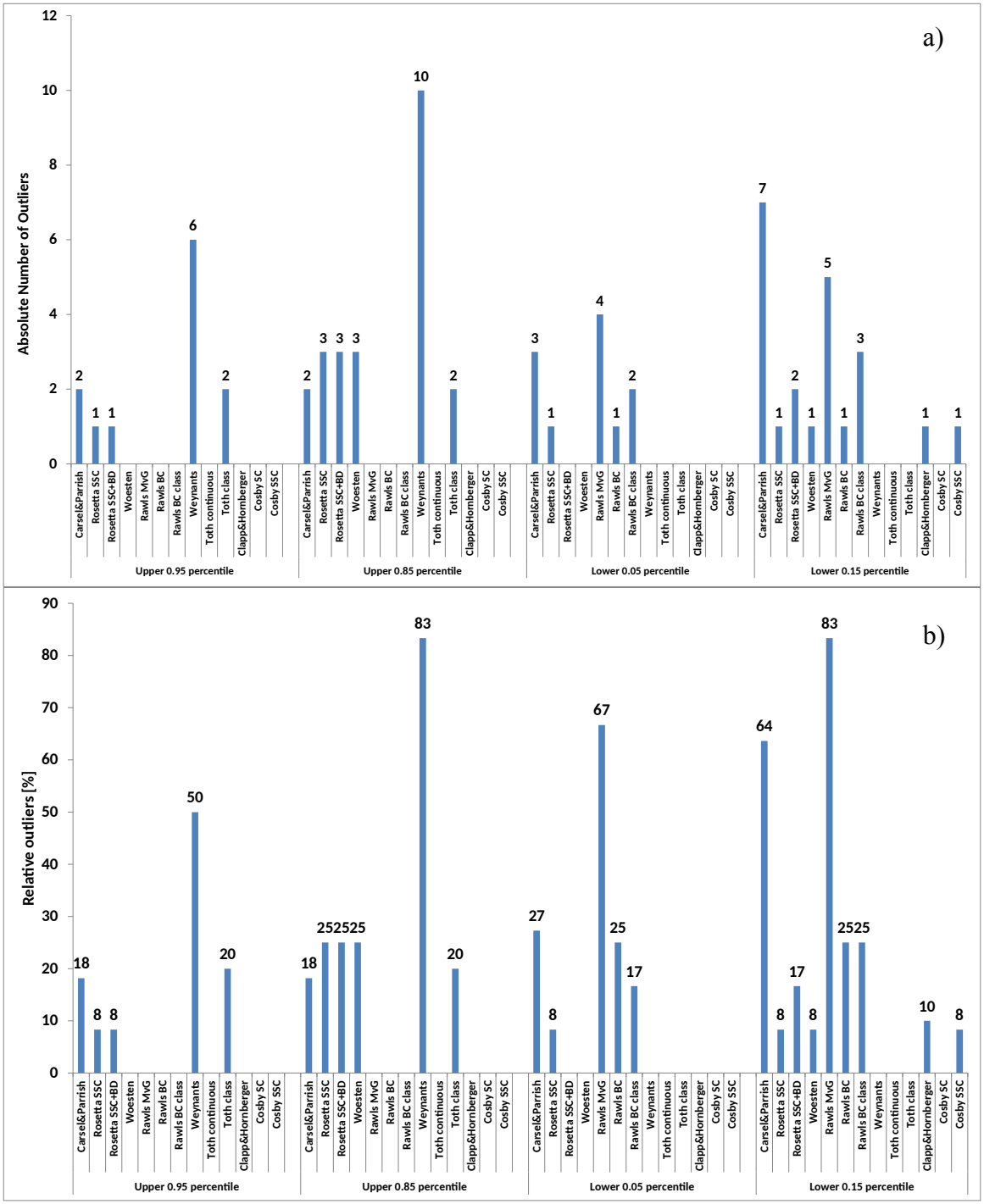


Annex Figure 4: Simulated cumulative actual evapotranspiration ET_a [cm] over the simulation period of 10988 days (30 years) for the homogeneous wheat vegetated scenario. Light and dark grey shaded area represent the 70 and 90 % tolerance interval, respectively.



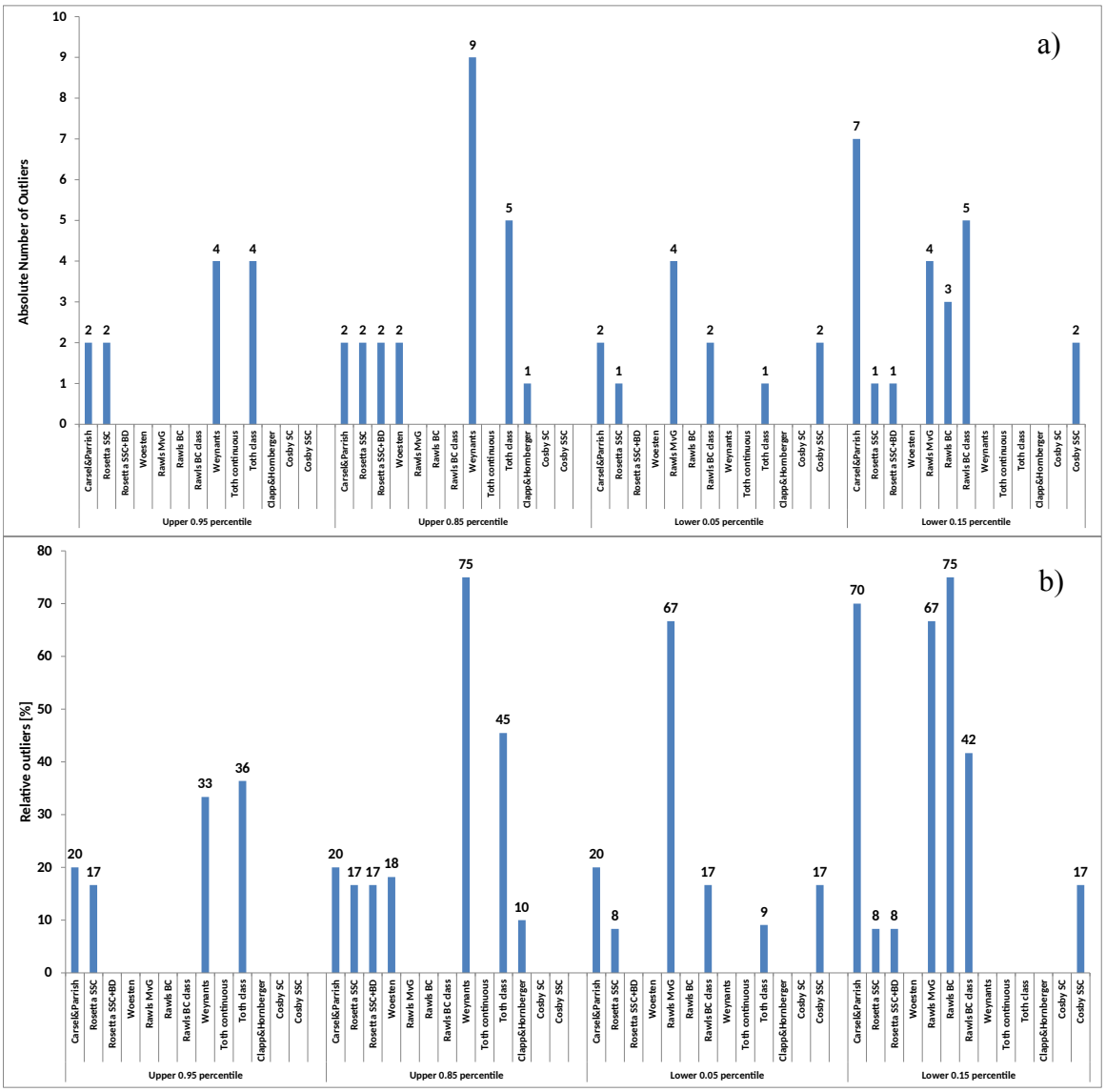
Annex Figure 5: Simulated cumulative actual evaporation E_a [cm] for two layered bare soil configurations (upper panel) and evapotranspiration ET_a [cm] for two layered grass configurations (second panel), for two layered wheat configurations (third panel), and for two layered wheat configurations with fluctuating ground water table (lower panel) over the simulation period of 10988 days (30 years) for the homogeneous wheat

vegetated configuration. Light and dark grey shaded area represent the 70 and 90 % tolerance interval, respectively. Note, that y-axis scales differ between configurations.



1480 **Annex Figure 6:** a) Absolute and b) relative number of outliers for simulated E_a/ET_a at t_{end} for the 12 USDA soil classes (11 for Clapp&Hornberger) and the homogeneous grass vegetated soil scenario.

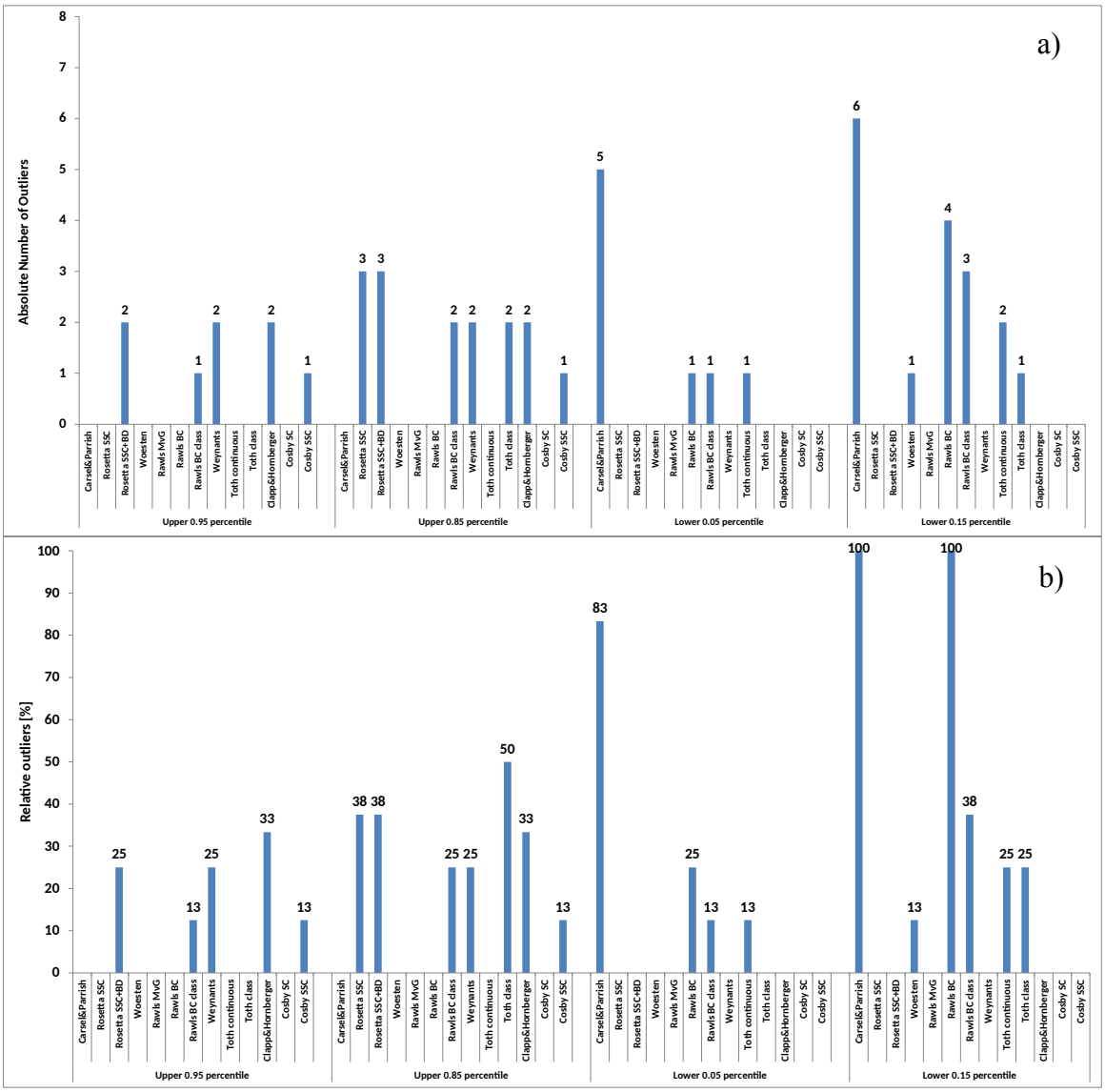
1485



1490 **Annex Figure 7:** a) Absolute and b) relative number of outliers for simulated E_a/ET_a at t_{end} for the 12 USDA soil classes (11 for Clapp&Hornberger) and the homogeneous wheat vegetated soil scenario.

1495

1500



Annex Figure 8: a) Absolute and b) relative number of outliers for simulated E_a/ET_a at t_{end} for the 12 USDA soil classes (11 for Clapp&Hornberger) and the layered scenarios (bare, wheat, grass, and wheat with fluctuating ground water table).