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 et al. (1984) 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

62   **1. Introduction**

63   Water fluxes and soil water content are key variables in the terrestrial system as they control

64   the exchange of water and energy between the land-surface and the atmosphere (e.g., Vereecken

65   et al., 2015). Modelling of the water flow in the unsaturated zone, and the uncertainty in the

66   parameters used to simulate water flow, has been a topic of intense research for many years,

67   both in the soil hydrological and land surface modelling community (Shao & Irannejad, 1999;

68   (Tietje & Tapkenhinrichs, 1993; Vereecken et al., 2008; Iwema et al. (2017)). Moreover,

69   climate modellers have studied the role of soil water content, and strongly related processes

70   such as evapotranspiration, in climate and atmospheric processes (Koster & Suarez 2001; Ek

71   & Holtslag, 2004; van den Hurk et al., 2008; Seneviratne et al., 2010; Groh et al., 2020). In this

72   context, we require reliable estimates of soil hydraulic properties at point to global scale

73   (Cornelis et al., 2001, van Looy et al., 2017). Measuring these properties is tedious, time and

74   cost expensive, and prone to measurement errors. Often, taking measurements is not feasible

75   due to the complexity and/or size of the terrestrial system under investigation. To overcome

76   this problem, pedotransfer functions (PTFs), which estimate these essential soil properties from

77   easily available soil parameters, such as soil texture, soil structure, bulk density, and soil organic

78   matter, have been developed. An extensive overview of existing PTFs was provided by

79   Vereecken et al. (2010) and by van Looy et al. (2017).

80   Soil properties used as basic input data to estimate the soil hydraulic properties with PTFs are

81   grouped into four categories: (1) soil particle size or soil texture, (2) easily measurable hydraulic

82   properties, (3) morphological properties, and (4) chemical properties (Espino et al., 1995,

83   Vereecken et al. 2010; van Looy et al. 2017; Rahmati et al., 2013; Neyshabouri et al., 2015;

84   Rahmati and Neyshabouri, 2016). In general, two different types of PTF can be distinguished,

85   namely point and parametric PTFs. Point PTFs estimate soil water content (or hydraulic

86   conductivity) values at predefined pressure head values (e.g., field capacity or wilting point),

87   whereas parametric PTFs provide the parameters than can be used in hydraulic functions (water

88   retention curve (WRC) and hydraulic conductivity curve (HCC)) of the Brooks & Corey (1964)

89   or van Genuchten (1980) formulation. The most useful PTFs developed in recent years are the

90   parametric PTFs because they can be used to calculate the WRC and HCC, which are used to

91   simulate the water fluxes in the numerical models. Secondly, PTFs are classified into class and

92   continuous PTFs. Class PTFs are look up tables, where the hydraulic parameters are listed for

93   typical soil textural classes (e.g., 12 USDA soil classes). Continuous PTFs, on the other hand,

94   use mathematical descriptions, e.g. regression functions, to calculate the hydraulic parameters

95   from the entire range of data inputs like e.g. soil texture.

96 It has been shown that PTFs are highly accurate for the area (or the input data range) they were

97 developed for, but have limited accuracy if applied outside these regions (Vereecken et al.,

98 2010). Several reviews about the accuracy and reliability of PTFs for the van Genuchten model

99 (VGM) have already been published (e.g., Wösten et al., 2001; Schaap, 2004; Donatelli et al.,

100 2004). Hereby, the predicted hydraulic function from the PTFs were compared to the measured

101 data and the goodness of fit of the prediction was evaluated. The authors used two metrics to

102 determine the performance of the PTF: 1) the term accuracy was related to the comparison

103 between predicted and measured values of water content or hydraulic conductivity that were

104 used to develop the PTF; 2) reliability was related to the evaluation of PTFs on measured values

105 that were different from those that were used to develop the PTFs (Wösten et al., 2001).

106 Reliability studies are typically validation studies such as those performed by Tietje &

107 Tapkenhinrichs (1993), Wösten et al. (2001), and Wagner et al. (2001). Despite much progress

108 in developing PTFs and in identifying appropriate PTF predictor candidates, some unresolved

109 or unexplained variability still exists at the level of the soil sample (Schaap & Leij, 1998), which

110 plays and important role when functional aspects of soils are being studied and analysed using

111 numerical models (e.g., Christiaens & Feyen, 2001). Functional aspects already studied are the

112 impact of PTFs on water supply capacity (Vereecken et al., 1992), ground water recharge

113 (Vereecken et al., 1992), and aeration (Wösten et al., 2001). In the study of Vereecken et al.

114 (1992), the authors showed that 90% of the variation in the predicted soil water supply was

115 attributed to estimation errors in hydraulic properties using the PTFs developed by Vereecken

116 et al. (1989, 1990). Chirico et al. (2010), on the other hand, evaluated the effect of PTF

117 prediction uncertainty on the components of the soil water balance at the hillslope scale. One

118 major result was that the simulated evaporation was much more affected by the PTF model

119 error than by errors resulting from uncertainty in the input data (e.g., soil texture).

120 Land surface models (LSMs), when embedded in numerical weather prediction or climate

121 models, generally operate at large scales (regional, continental to global scales) and rely on

122 PTFs to predict the hydraulic functions needed to solve the Richards equation for the water

123 flow. Different LSMs use different PTFs for this purpose (Vereecken et al., 2019). As

124 Vereecken et al. (2019) showed, not only different PTFs but also different hydraulic models

125 (Campbell, Brooks and Corey, or Mualem-van Genuchten) are in use. Knowing that different

126 PTFs and/or the choice of the hydraulic model will impact the outcome of the water flow

127 simulations (e.g., Gruber et al., 2006; Yakirevich et al., 2013), a key question is how recently

128 launched LSM inter-comparison activities of the Land Surface Schemes (LSSs) embedded in

129 LSMs, such as those under the Global Energy and Water Exchanges (GEWEX) GLASS project

130 (https://www.gewex.org/panels/global-landatmosphere-system-study-panel/) or model inter-

131 comparisons initiated by the World Climate Research Programme (https://www.wcrp-

132 climate.org/wgcm-cmip/wgcm-cmip6, of which GEWEX is part), such as CMIP6 and its

133 predecessors, will be impacted by the choice of PTF.

134 Therefore, the aim of this study is to systematically analyse the functional sensitivity to the

135 choice of different PTFs using a physically-based numerical model. As the 'truth' of this model

136 exercise is unknown, the performance of the model runs with its hydraulic parameters derived

137 from a set of individual PTFs will be evaluated against the ensemble mean as best predictor, as

138 well as against the 70 and 90 tolerance intervals of the ensemble range. The numerical exercise

139 is structured in the following way: 1) model runs for homogeneous soil profiles without

140 vegetation, 2) homogeneous soil profile covered with grass and wheat, 3) layered bare soil, 4)

141 layered vegetated soil (grass and wheat), and 5) influence of a fluctuating water table in a

142 layered grass vegetated soil. Finally, additional soil physical properties were calculated based

143 on the estimated soil hydraulic parameters obtain from the PTFs, which were used to explain

144 the differences observed in simulated water fluxes. As some LSMs also use class PTFs (van

145 Looy et al., 2018, Vereecken et al., 2019), we will also analyse the use of this type of PTF, and

146 the associated errors when simulating water fluxes. We formulate three hypotheses 1) the use

147 of different PTFs will lead to systematically different hydrological states and fluxes (e.g., net

148 infiltration, evapotranspiration, root zone water availability, drainage), 2) some PTFs can be

149 identified which perform distinctively differently from the ensemble spread in terms of 90 %

150 tolerance interval outliers, and 3) the differences in predicted states and fluxes simulated with

151 inputs from different PTFs will be reduced with increasing model setup complexity.

152

153 **2. Materials and Methods**

154 *2.1. Hydraulic functions*

155 Three pairs of hydraulic functions are widely used in hydrological modelling, namely those

156 developed by Brooks and Corey (1964), Campbell (1974), and van Genuchten (1980.

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

158
$$S_e = \begin{cases} |\alpha h|^{-n} & h < -1/\alpha \\ 1 & h \geq -1/\alpha \end{cases} \qquad [1]$$

159 where α is the reciprocal of the air entry value (or bubbling pressure) [cm^{-1}], n is a dimensionless

160 shape parameter [-] (related to $1/b$ for the original Brooks-Corey b parameter), h is the pressure

161 head [cm], and S_e is the effective saturation [-] given by:

162
$$S_e = \frac{\theta - \theta_r}{\theta_s - \theta_r} \qquad [2]$$

163   where θ is the actual water content [cm^3 cm^{-3}], θ_r is the residual water content [cm^3 cm^{-3}], and

164   θ_s is the saturated water content [cm^3 cm^{-3}].

165   The unsaturated hydraulic conductivity is given by:

166    $$K = K_s \, S_e^{2/n+3} \qquad\qquad [3]$$

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

168   The Campbell (1974) water retention function is a modification of that introduced by Brooks

169   and Corey (1964), with θ_r set to 0.

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

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

172   where m is a shape factor related to n by $m = 1-1/n$.

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

174    $$K_{(h)} = K_s S_e^{\lambda} \left[1 - \left(1 - S_e^{1/m}\right)^m\right]^2 \qquad\qquad [5]$$

175   where λ is the tortuosity factor [-].

176   *2.2. Pedotransfer Functions*

177   For this study 13 pedotransfer functions (PTFs) were used, whereby eight predict the hydraulic

178   parameters for the MvG (van Genuchten, 1980) and five PTFs predict the parameters in the BC

179   (Brooks & Corey, 1964) or Campbell (Campbell, 1974) functions. Out of these 13 PTFs, four

180   were so called *class-transfer* functions, where only the USDA textural classes will be used as

181   input for the prediction of the hydraulic parameters. It has to be noted that the PTF of Clapp

182   and Hornberger, (1978) (from here on Clapp&Hornberger) does not specify hydraulic

183   parameters for the silt class. All other PTFs use textural information (gravimetric percentage of

184   sand, silt, and clay) as basic inputs. Additionally, some PTFs require information about bulk

185   density (BD) such as the PTFs of Schaap et al. (2001) (here referred to as Rosetta SSC+BD),

186   Wösten et al. (1999) (here Woesten), Weynants et al. (2009) and Weihermüller et al. (2017)

187   (here Weynants), and that of Tóth et al., (2015) for the topsoil (here 'Toth continuous'). Others

188   need information about the organic carbon content (C_{org}), which are the Woesten, Weynants,

189   and 'Toth continuous' PTFs.

190   Soil organic carbon is used as a predictor the PTFs as it affects soil bulk density, hydraulic

191   conductivity, and water retention because of its effect on soil structure and adsorption properties

192   (van Genuchten & Pachepsky, 2011).

193   Total porosity ϕ, as estimated from bulk density (BD), is used only by Rawls and Brakensiek

194   (1985) (here Rawls MvG), and pH and cation exchange capacity (CEC) are inputs in 'Toth

195   continuous'. An overview of all used PTFs, their abbreviations, and their inputs is provided in

196   Tab. 1. The region from where data were taken to train the PTF are from either the USA or

197   Europe. Rosetta is the only PTF combining two data regions, whereas Weynants PTF is based

198   on samples from Belgium only. In addition, the number of samples used for PTF development

199   greatly differs, ranging from 5320 for Rawls PTFs to 166 for Weynants. Important for the PTF

200   development is the data used to generate the PTFs, whereby either only retention data ($\theta_{(h)}$) or

201   a combination of retention ($\theta_{(h)}$) and unsaturated hydraulic conductivity ($K_{(h)}$) data was used.

202     $\theta_{(h)}$ and $K_{(h)}$ were used in the development of the PTFs Rosetta, Woesten, Weynants, and Toth,

203   whereby the percentage of available $K_{(h)}$ data is typically low compared to the availability of

204     $\theta_{(h)}$ data, generally due to the more complex and laborious procedures required to determine

205     $K_{(h)}$. Even though in some cases both types of data ($\theta_{(h)}$ and $K_{(h)}$) were used in the development

206   of some PTFs, the data were not jointly inverted to estimate the hydraulic parameters, meaning

207   that Rosetta, Woesten, and Toth fitted the hydraulic parameters sorely on the retention curve

208   and used the fitted α and n values of the Mulaem van Genuchten equation to predict $K_{(h)}$. In

209   contrast, Weynants used joint inversion of both hydraulic characteristics ($\theta_{(h)}$ and $K_{(h)}$)

210   simultaneously to estimate the parameters including a near saturation hydraulic conductivity

211   K_s^* at a predefined pressure head of – 6 cm. All other PTFs either used the closed form

212   expression of van Genuchten (1980) or Brooks and Corey (1964) to predict $K_{(h)}$, using the

213   estimated parameters from the retention data, together with measured K_s values, to estimate the

214   unsaturated hydraulic conductivity based on either van Genuchten (1980) or Brooks and Corey

215   (1964).

216   In this study, we will compare model simulations for 12 soil textural classes. For the estimation

217   of hydraulic parameters from texture based continuous PTFs a representative soil texture was

218   used for each soil class located in the centre of the respective class area in the textural triangle;

219   bulk density and C_{org} were set to 1.4 g cm^{-3} and 1 %, respectively. The texture of the

220   corresponding class is depicted in Fig. 1 and the predicted hydraulic parameters for all applied

221   PTFs are listed in Annex Tab. 1 and Annex Tab. 2.

222   In general, it is known that relatively small changes in the shape of the soil water retention

223   curve near saturation can significantly affect the results of numerical simulations of water flow

224   for variably saturated soils, including the performance of the numerical stability and rate of

225   convergence (Vogel et al., 2001; Schaap & van Genuchten, 2006). To address this problem,

226   especially in fine textured soils, the estimated air entry value (i.e., the reciprocal of α) from the

227 PTF for the van Genuchten formulation (Eq. 4) was set to -2 cm as proposed by Vogel et al.

228 (2001), whenever the originally proposed set of hydraulic properties from the PTF did not lead

229 to numerical convergence.

230   *2.3. Numerical Modelling*

231 For the simulation of vertical water flow, the one-dimensional Richards equation (Eq. 6) was

232 solved using the finite element code HYDRUS-1D (Šimůnek et al., 2008; Šimůnek & van

233 Genuchten, 2008):

234

235 $$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[K_{(\theta)}\left(\frac{\partial h}{\partial z} + 1\right)\right] - Q \qquad [6]$$
236

237 where z represents the vertical coordinate [cm], positive in the downward direction, Q is the

238 source/sink term, θ is the volumetric water content [$cm^3\ cm^{-3}$], and $K_{(\theta)}$ is the unsaturated

239 hydraulic conductivity [$cm\ d^{-1}$] as a function of water content.

240 A 200 cm soil profile was simulated, and the lower boundary condition of the flow domain was

241 defined as free drainage, which is typically used when the ground water table is far below the

242 soil surface. As a second option, a fluctuating ground water table was prescribed as a Dirichlet

243 boundary condition. For the variable ground water table depth, a simple sine curve of the ground

244 water table fluctuations was generated by using Eq. [7] for the 10988 days of climatic data (see

245 end of this section for a description of the model driving data):

246

247 $$y(t) = A_s sin(2\pi f t) \qquad [7]$$

248

249 where A_s is the amplitude of the sine curve, which is defined as the maximum displacement of

250 the function from its centre position, and shows the height of the curve (fluctuation) (here, $A_s =$

251 100 cm), 2π is the natural period of the sine curve, f is the frequency (here, $f = 1/365\ days^{-1}$),

252 and t is the time period of the sine curve (here, $t = 1$ to 10988 days).

253 The upper boundary condition in HYDRUS was set to an atmospheric boundary with surface

254 run off. The domain was non-equally discretized with 401 nodes, with finer discretization at

255 the top to account for the stronger flow dynamics close to the soil surface. Pressure head was

256 used for initialization of the soil profile with linearly decreasing potentials between the bottom

257 (0 cm) and the top (200 cm) node (i.e., hydrostatic equilibrium).

258 For the simulations, different setups were chosen with varying complexity. A simple

259 homogeneous soil profile without vegetation was selected as the simplest case, to study the

260 impact of the choice of different PTFs. Complexity was increased by adding different

261 vegetation covers (grass and wheat). In both cases, growth was not simulated and both crops

262 covered the soil throughout the entire year. Potential evapotranspiration, ET_0, was split into soil

263 evaporation, E, and transpiration, T, by setting T to 75 % of ET_0. Also, rooting depth was

264 assumed to be the same (0-30 cm) for both vegetation covers with a linear decrease in root

265 density from the top soil layer to the maximum rooting depth. The root water uptake reduction

266 model of Feddes et al. (1978) was used, based on the parameter values of Wesseling (1991) for

267 both grass and wheat vegetation, as taken from the HYDRUS embedded look up table (see

268 Tab. 2). Therefore, the only difference between both vegetation scenarios was the root water

269 uptake. This simplification in terms of growing season and rooting depth was done to simplify

270 the comparison of the simulation results, by ensuring that root water uptake will not be from

271 different soil layers when grass is replaced by wheat. In a next step of increasing complexity,

272 soil layering was introduced, whereby two layering schemes were assumed. 1) Sandy loam over

273 silt loam overlaying a loamy sand and 2) silt loam over silty clay loam overlaying a silty clay,

274 respectively. For the layered profiles the first layer was set to extend from 0 to 50 cm, the second

275 layer from 50 – 100 cm, whereas the third layer occupied the rest of the profile (100 – 200 cm).

276 Again, the same vegetation parameters as for the homogeneous soil were used. Finally, the

277 layered system covered by wheat with a fluctuating groundwater table was simulated. Figure 2

278 shows a schematic view of the seven model scenarios used in this study.

279 Thirty years (10988 days) of daily climatic data (comprising precipitation and Penmen-

280 Monteith potential ET) from 1982 to 2011 were taken from North Rhine-Westphalia, Germany

281 (mean NRW climatic data) as used by Hoffmann et al. (2016) and Kuhnert et al. (2017). The

282 climate is humid temperate.

283   *2.4. Statistics and data evaluation*

284 The HYDRUS-1D outputs were used to analyse the fluxes of the soil water balance, i.e. actual

285 evaporation (E_a), actual evapotranspiration (ET_a), drainage (D), and run off (R). For the data of

286 the 13 different PTFs the cumulative flux was selected and the arithmetic mean, 0.05 and 0.95

287 percentiles (which spans the 90 percent tolerance interval), and the 0.15 and 0.85 percentiles

288 (which spans the 70 percent tolerance interval), were calculated for the last data point of the

289 cumulative fluxes (t_{end} = day 10988) for each soil textural class. In a next step, outliers based

290 on the percentiles were calculated and flagged according to Eq. [8] and [9]:

291

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

293 *If value of a PTF for $flux(x) < 0.05$ percentile, value = −1, else value = 0*   [9]

294

295   where *flux(x)* is the cumulated flux (e.g., actual evapotranspiration) at t_{end} for each individual

296   PTF used to simulate the soil class for each model scenario. In other words, if the flux value,

297   *flux(x)*, at t_{end} exceeds (is less than) the percentile span calculated, the simulation was flagged

298   with a 1 (-1), whereas if the flux value, *flux(x)*, at t_{end} lies within the given percentile, the

299   simulation was flagged with a 0. The same procedure as for the 90 percent tolerance interval

300   was repeated for the 70 percent tolerance interval. In order to present variability of the hydraulic

301   parameters or simulated fluxes, the coefficient of variation (CV) was calculated, relating the

302   standard deviation to the mean value. The CV was expressed as a percentage.

303   For the analysis of differences in soil hydraulic properties estimated by the PTFs, the

304   comparison of two group means was performed with the non-parametric Wilcoxon rank sum

305   test using Matlab® *ranksum* function using the probability $p = 0.05$. Principal component

306   analysis (PCA) was performed, also with Matlab®.

307   To help interpret the differences between model runs with regards to the water fluxes, the matric

308   flux potential, *MFP*, the characteristic length, L_C, the sorptivity, *S*, and characteristic time t_{grav}

309   was calculated as explained next.

310

311   The matric flux potential *MFP* [cm^2 d^{-1}] is a convenient bulk soil hydraulic property that is

312   often used in soil water movement studies (e.g., Raats, 1977; Pullan, 1990; Grant & Groenevelt,

313   2015), which is defined as the integral of the hydraulic conductivity $K_{(h)}$ [cm d^{-1}] over the

314   pressure head *h* [cm] starting at an arbitrary reference pressure head h_{ref}:

315

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

317

318   where h_{ref} was set to permanent wilting point at $h = -15{,}000$ cm according to Pinheiro et al.

319   (2018) and *h* set to full saturation ($h = 0$).

320

321   As a second soil physical feature, the characteristic length of bare soil evaporation, L_C [cm],

322   was calculated. According to Lehmann et al. (2008 and 2018), L_C is the maximal extent of the

323   capillary flow region to supply water to evaporating surface. L_C is determined by the range of

324   capillary pressure between large and small pores driving the capillary flux against gravity

325   (expressed as head difference, denoted as gravity length L_G) and the hydraulic conductivity K_{eff}

326   of the supply region. Formally, L_C is defined via:

327

328 $$L_C = \frac{L_G}{1 + E_0/K_{eff}} = \frac{\frac{(1-m)}{\alpha}\left(1+\frac{1}{m}\right)^{(1+m)}}{1 + E_0/4K(h_{crit})}$$ [11]

329

330 where α and m are the van Genuchten parameters used in Eq. [4] and E_0 is potential evaporation

331 rate. To calculate effective conductivity, K_{eff} was estimated as $4K_{crit}$ (Haghighi et al., 2013;

332 Lehmann et al., 2018), whereby K_{crit} is the unsaturated hydraulic conductivity at critical water

333 content and capillary pressure when capillary pathways start to disconnect (Lehmann et al.,

334 2008). The critical capillary pressure and gravity length are determined based on linearization

335 of the soil water retention curve. For the van Genuchten formulation used in Eq. [11], the

336 linearized retention curve consists of the tangent to the inflection point, and L_G and h_{crit} can be

337 expressed analytically. For Brooks and Corey, the values are determined numerically with a

338 line passing through the air-entry value ($S_e = 1$, $h = h_b$) and a particular point on the retention

339 curve that is closest to ($S_e = 0$, $h = h_b$).

340

341 The soil sorptivity, S [cm d$^{0.5}$], which is defined as a measure of the capacity of a porous

342 medium to absorb or desorb liquid by capillarity, was calculated assuming a soil column with

343 uniform initial water content and infinite length, following the approach of Parlange (1975) as

344 described in Moret-Fernández et al. (2017), Lotorre et al. (2018), and Rahmati et al. (2019):

345

346 $$S^2(\theta_s, \theta_i) = \int_{\theta_i}^{\theta_s} D(\theta)[\theta_s + \theta - 2\theta_i]d\theta$$ [12]

347

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

349

350 $$D(\theta) = K(\theta)\frac{dh}{d\theta}$$ [13]

351

352 For the initial water content, θ_i, the maximum reported residual water content (θ_r) for all MvG

353 parameters used in this study (0.192 cm^3 cm^{-3}) was used for all PTFs based on MvG and 0.12

354 cm^3 cm^{-3} for all PTFs based on BC.

355

356 Finally, the so-called characteristic time (t_{grav}) was calculated according to Philip (1957), which

357 determines the 'time' t where gravitational forces become dominant ($t \gg t_{grav}$), while for $t \ll$

358 t_{grav} capillary forces remain dominant over gravitational forces (Rahmati et al., 2020):

359

$$360 \quad t_{grav} = \left(\frac{S}{K_s}\right)^2 \qquad [14]$$

361

362 Two final related soil properties that were calculated were the characteristic time for the

363 attainment of field capacity FC, τ_{FC} (d), and the elapsed time required for attainment of FC,

364 denoted as t_{FC} (d]), see Assouline and Or (2014). To do so, the effective soil saturation at field

365 capacity S_{FC} [Eq. 15] and the unsaturated hydraulic conductivity at field capacity $K(S_{FC})$ [Eq.

366 16] were calculated by:

367

$$368 \quad S_{FC} = \left[1 + \left\{\left(\frac{n-1}{n}\right)^{(1-2n)}\right\}\right]^{\left(\frac{1-n}{n}\right)} \qquad [15]$$

369

$$370 \quad K_{FC} = K_s S_{FC}^{0.5} \left[1 - \left(1 - S_{FC}^{(1/m)}\right)^m\right]^2 \qquad [16]$$

371

372 Moreover, the quantity of drainable water from the soil profile to depth z at field capacity FC,

373 Q_{FC}, expressed as equivalent water depth (dimensions of length), has to be calculated as:

374

$$375 \quad Q_{FC} = z(\theta_s - \theta_r)(1 - S_{FC}) \qquad [17]$$

376

377 Here, z was set to 30 cm to match the maximal rooting depth for convenience, as z only scales

378 with total Q_{FC}.

379 Consequently, a characteristic time [d] for the attainment of FC, τ_{FC}, can be deduced from the

380 ratio of these two quantities, Q_{FC} and K_{FC}:

381

$$382 \quad \tau_{FC} = \frac{Q_{FC}}{K_{FC}} \qquad [18]$$

383

384 The drainage dynamics can now be linked to the elapsed time [d] required for attainment of FC,

385 denoted as t_{FC}:

386

$$387 \quad t_{FC} = \frac{z(\theta_s - \theta_r)}{K_m} ln\left(\frac{K(S_{FC})}{Ks}\right) \qquad [19]$$

388

389 where K_m is the effective hydraulic conductivity that represents the mean value of $K(S_{FC})$

390 weighted by S_{FC} (representing the available relative cross section of flow):

391 $$K_m = \frac{\int_0^1 S_{FC}K(S_{FC})dS_{FC}}{\int_0^1 S_{FC}dS_{FC}}$$ [20]

392

393 **3. Results and Discussion**

394 *3.1.Predicted hydraulic parameters and hydraulic functions*

395 In a first step, the retention and hydraulic conductivity curves for the 13 PTFs and the 12 USDA

396 soil classes were plotted based on the estimated soil hydraulic parameters listed in Annex Tab.

397 1 and 2. As an example, the retention and hydraulic conductivity curves for the USDA textural

398 class sand is plotted in Fig. 3 (see Annex Fig. 1 for all retention and hydraulic conductivity

399 curves of all USDA textural classes). Figure 3 shows that the retention curves based on the 13

400 PTFs greatly differ along the entire pressure head range. For these curves, the saturated water

401 content, θ_s, varies between 0.472 for Rawls MvG and 0.375 cm^3 cm^{-3} for Cosby SSC with a

402 mean of 0.417 cm^3 cm^{-3} over all PTFs. The corresponding coefficient of variation (CV) is 7.5

403 % for the sandy soil. The residual water content, θ_r, varies between 0 for those PTFs setting θ_r

404 to 0 such as Weynants, Clapp&Hornberger, Cosby SC, and Cosby SSC, to 0.061 cm^3 cm^{-3} for

405 the 'Toth class' PTF with a mean of 0.029 cm^3 cm^{-3} and a CV of 80.1%, indicating a much

406 higher variability in terms of CV in θ_r compared to θ_s. An even larger variability can be found

407 in the saturated hydraulic conductivity, K_s, for which we found a maximum value of 1520.6 cm

408 d^{-1} for Clapp&Hornberger and a minimum of 8.3 cm d^{-1} for the 'Toth class' with a mean of

409 315.8 cm d^{-1} (CV = 129.2 %). In general, the smallest CV values (data not shown) for all USDA

410 textural classes were found for θ_s, with lowest ranging from 2.4% for the silty clay loam class

411 to CV = 7.5 % for the sand class. (Larger variability was observed in θ_r with the lowest

412 coefficient of variation in the loamy sand class (CV = 76.3 %) and highest in the silty clay loam

413 class (CV = 106.6 %). As expected, K_s showed the largest variability, with lowest CV in the

414 loam class (91.7 %) and largest in the clay loam class (215.2 %). The larger CV values for the

415 K_s estimation is not surprising as a large uncertainty in predicted K_s has been already widely

416 reported (e.g., Jaynes & Tyler, 1984; Ahuja et al., 1985; Tietje & Hennings, 1996; Schaap et

417 al., 1998). Furthermore, it has to be noted that the PTF developed by Weynants et al. (2009)

418 predicts K_s^* instead of K_s, where K_s^* is a hydraulic conductivity acting as a matching point at

419 suction head h = -6. Therefore, some slightly lower K_s (here K_s^*) value will be predicted by

420 Weynants' PTF. On the other hand, there seems to be a clear grouping among the class PTFs,

421 with regards to the estimation of K_s. Clapp&Hornberger predicted the highest values for six

422 classes (sand, loamy sand, sandy loam, loam, silt loam, and silty clay loam), followed by the

423 PDF of Woesten for five soil classes (sandy clay loam, clay loam, sandy loam, silt loam, and

424 clay). Even more pronounced is the picture for the prediction of lowest K_s, whereby the PTF of

425 Rawls MvG estimated the smallest K_s values for 11 soil textural classes, except for sand,

426 whereas the 'Toth class' PTF showed the lowest K_s values. Unfortunately, the α and n (or $1/b$)

427 values cannot be directly compared between the BC and MvG approaches, as both parameters

428 have a slightly different physical meaning.

429 *3.2.Numerical model performance*

430 As numerical stability of the simulation is one of the crucial aspects in the choice and

431 application of the PTF, especially for large scale modelling, we analysed each PTF with respect

432 to numerical convergence, when using HYDRUS. For each PTF and the seven model scenarios

433 (see Fig. 2), 44 individual model runs were performed: for each PTF, the three homogeneous

434 soil layer model scenarios were modelled for each soil textural class (these are 36 model runs).

435 In addition, the four layered configurations are run for a coarse and a fine soil layering, resulting

436 in eight model runs; hence, a total of 44 model runs per PTF were obtained. Note that for the

437 Clapp and Hornberger (1978) PTF, only 41 model runs were performed as no parameters were

438 reported for the silt class. For 486 model runs, out of the total 569 model runs (i.e., 85 %),

439 convergence was achieved. A total of 184 out of 217 (85%) of the model runs for the BC and

440 279 out of 352 (79 %) for the MvG parameterization converged, even though it has been

441 reported that the BC type function sometimes prevents rapid convergence and might therefore

442 cause numerical problems. This was deemed to be caused by the discontinuity present in the

443 slope of both the soil water retention and unsaturated hydraulic conductivity curves (van

444 Genuchten, 1980).

445 For those cases where the simulation did not converge for the MvG parameters, the air entrance

446 value (the inverse of α) was set to -2 cm, and the model was rerun. This procedure increased

447 the total number of converged MvG simulation runs to 302 (86 %), which is a similar

448 percentage to that obtained for BC. Looking at individual PTFs (see Tab. 3), we can see that

449 the Rosetta SSC+BD and Cosby SSC seem to be numerically very stable with 100 % converged

450 runs. The Woesten and 'Toth continuous' function converged for > 95 % of the runs, after

451 setting the $1/\alpha$ to -2 cm. On the other hand, the lowest convergence was found for the Rawls

452 MvG and Rawls BC with 43 and 39 %. Unfortunately, using $1/\alpha$ = -2 cm did not improve

453 convergence for Rawls MvG.

454 The reason why some PTFs prohibited the HYDRUS model from converging is quite apparent

455 for some cases. For example, Rawls MvG and Rawls BC yielded very low K_s values of 0.8 cm

456   day^{-1} for the loam and sandy clay class and ≤ 0.3 cm day^{-1} for clay loam, silt, and silt loam

457   class. Extremely low values were obtained for silty clay loam (0.04 cm day^{-1}) and clay (0.004

458   cm day^{-1}), almost allowing no infiltration at all. Another extremely low K_s value was predicted

459   by the 'Toth class' PTF for silty clay, with 0.01 cm day^{-1}; these unrealistically low values again

460   led to numerical instabilities.

461   It has to be noted that the reported convergence here is only valid for the numerical model

462   (HYDRUS-1D) used in this exercise with the given numerical (convergence) default criteria,

463   vertical discretization and temporal resolution, and atmospheric boundary conditions. The

464   performance of these PTFs may change if a different numerical scheme, e.g. solving the

465   Richards equation in the mixed or diffusivity form were used, or a different spatial

466   discretization and/or temporal resolution. Furthermore, the lack of certain processes in our

467   simulations (e.g. coupled heat and water transport or evaporation from the wet canopy), or the

468   nature of the atmospheric forcings (e.g. a difference in rainfall frequency and amount) will

469   affect the likelihood of convergence.

470

471   *3.3.Fluxes and outliers*

472   *3.3.1   Simulated fluxes over time*

473   Firstly, the simulated cumulative fluxes were analysed. ET_a (vegetated surface) or E_a (bare soil)

474   is a key flux as it indirectly contains information of the net infiltration into the soil profile (net

475   daily infiltration = daily sum of precipitation – daily sum of ET_a or E_a), deep drainage (over

476   long-run), and plant available water in the root zone. Furthermore, ET_a or E_a determines the

477   return of water from the soil profile to the atmosphere, and as such affects the land surface

478   energy budget. Cumulative ET_a or E_a data for each scenario/soil class combination was plotted

479   and the arithmetic mean of all data (model ensemble mean, MEM) for each combination, as

480   well as the spread of the data, was calculated by the 70 and 90 percent tolerance interval

481   according to Eq. [8] and [9].

482   As an example of the high variability in simulated fluxes, the simulated cumulative E_a, for the

483   homogeneous bare soil scenario of *loamy sand* texture, over the entire simulation period of 30

484   years, that ends on day 10988 (t_{end}), is plotted in Fig. 4a. There is a large variability between

485   the various simulations based on the 13 PTFs. MEM at t_{end} is 1692 cm (564 mm year^{-1}). The

486   smallest simulated cumulative E_a was 1273 cm (424 mm year^{-1}) for the Carsel&Parrish PTF,

487   and largest, with 2043 cm (681 mm year^{-1}), for Weynants. The difference of 257 mm year^{-1}

488   between the largest and smallest simulated E_a, and their deviation of 140 and 117 mm year^{-1}

489   from MEM clearly indicates that the choice of PTF substantially affects the estimation of the

490 E_a for this soil class. In contrast, low variability was found for cumulative E_a for the bare

491 homogeneous *clay loam* (see Fig. 4b). Notably, two out of the 13 simulations did not converge

492 (Rawls MvG and Rawls BC), which potentially also impacts the variability in simulated fluxes.

493 Nevertheless, for the remaining 11 simulations the lowest simulated flux was 1744 cm (581

494 mm year^{-1}) for Rawls class PTF and largest for Weynants PTF with 2041 cm (680 mm year^{-1})

495 (for this soil, MEM = 1893 cm or 631 mm year^{-1}). Overall, the difference between the largest

496 and smallest flux is only 99 mm year^{-1}, i.e., 2.5 times smaller than the difference found for the

497 loamy sand. As E_a will be also be influenced by the precipitation entering the soil (total

498 precipitation over 30 years = 2479.7 cm (827 mm yr^{-1})), we also looked at the cumulative

499 runoff. For most soil textural class/PTF combinations, still for the homogeneous bare soil

500 scenarios, runoff is low or negligible, with zero runoff, or values < 1 cm over 30 years, for 121

501 model runs, which is equivalent to 88 % of runs. Nine simulations (7 %) returned a runoff >1

502 cm but <10 cm, and eight exceeded 10 cm over 30 years (7 %). The highest cumulative runoff

503 was generated for the Rawls MvG/silt combination with 675 cm (27 % of total precipitation)

504 followed by Rawls silt loam with 664 cm and Rawls MvG loam with 389 cm. These three

505 combinations have also been flagged as outliers of the lower 0.15 percentile for total cumulative

506 evaporation at t_{end}, which can be explained by the fact that less water enters the soil, and

507 therefore less will be evaporated. The same holds for the Carsel&Parrish PTF and silty clay

508 loam (runoff = 135.1 cm) and sandy clay (runoff = 70 cm) combinations. On the other hand,

509 the combination 'Toth class' PTF/silt loam generated 122.9 cm runoff but was classified as an

510 upper 0.95 percentile outlier, generating more E_a. Finally, Rawls BC/silt and 'Toth

511 continuous'/silty clay combinations generated run off of 156.1 and 43.7 cm, respectively, yet

512 are not classified as outliers. In general, runoff generation is linked to low K_s values (see Annex

513 Tab. 1 and 2). An overview of all cumulative E_a fluxes at t_{end} for the bare soil scenarios is

514 plotted in Fig. Annex 2.

515 Our findings with regards to ET_a for the vegetated scenarios (grass and wheat), still with a

516 homogeneous soil profile (Annex Fig 3 to 4), were comparable. For some soil classes, such as

517 clay, clay loam, and silty clay loam, variability between PTFs was low, for both grass and

518 wheat, whereas the ET_a for the sandy and sandy loam soils showed consistently high variability.

519 In contrast, ET_a for the loamy sand class exhibited relatively high variability for grass (as was

520 also the case for bare soil) and a slightly smaller one for the wheat scenario configuration. For

521 the other soil textural classes, the picture is less clear. Again, as was the case for the bare soil,

522 there is a substantial number of soil class/PTF combinations that result in runoff. A slightly

523 larger, compared to the bare soil scenario, percentage of simulations with run off > 1 cm was

524   found for the grass (18 %) and the wheat (20 %) scenarios. Moreover, maximum runoff at t_{end}

525   value increased from bare (675 cm for Rawls MvG silt) via grass (859.2 for Rawls MvG

526   silty clay) to the wheat scenario (999.2 cm for Rawls BC silty clay). Surprisingly, eight out of

527   twelve soil class/PTF combinations yielding runoff > 100 cm were not flagged as outliers for

528   the 90 % tolerance interval for the grass and four out of 10 for the wheat. There are some

529   unexpected findings, namely that the simulation for the Toth continuous' PTF yielded 46.1 and

530   11.7 cm runoff, respectively, for the silty clay and silty clay loam under wheat vegetation,

531   despite the fact that the ET_a flux at t_{end} was flagged as an outlier of the upper 0.95 percentile,

532   indicating relatively high evaporation with respect to the model ensemble.

533   Finally, the simulation for the scenario of sandy loam overlying silt loam and loamy sand

534   plotted in Annex Fig 5 showed much lower variability for E_a compared to E_a of the homogenous

535   profile with the texture of the uppermost layer (silt loam in Annex Fig. 2). This indicates that

536   soil layering will reduce the effect of the choice of PTF on the cumulative evaporation. This

537   holds true even more for the layered bare soil scenario where silt loam overlies silty clay loam

538   that is overlying silty clay. Again, the variability in E_a for the layered system is much lower

539   than that of the homogeneous silt loam, that forms the first layer in the vertically heterogeneous

540   soil profile. Besides, it can clearly be seen that when vegetation is introduced, variability

541   increases slightly, which is reflected in the coefficient of variation (CV) of the flux at t_{end}, where

542   for the layered profile topped by sandy loam the CV increased from 3.9, via 5.1 to 4.9 % for

543   the bare, grass, and wheat vegetation scenario. For the profile with the first layer consisting of

544   silt loam, CV values were 0.5, 3.7, and 3.7 % for the bare, grass, and wheat vegetation,

545   respectively. Introducing a fluctuating ground water table increased the CV substantially to 13.6

546   % for both vegetated layered systems. Variability in simulated E_a or ET_a for the layered

547   scenarios can partly be explained by a large reduction in runoff. In total only two simulations

548   (2 %) for the Carsel&Parrish (sandy loam topped layered profile under grass vegetation (290.2

549   cm) and sandy loam topped layered profile for the wheat vegetation and ground water

550   fluctuation (302.6 cm)) exceeded run off of 100 cm. A further four exceeded the run off

551   threshold of 1 cm (3 combinations for Carsel&Parrish and one for Rawls BC).

552   Overall, the choice of PTF substantially affects the simulated values of E_a or ET_a for most soil

553   classes, irrespective of the fact whether the soil was bare, where the water (vapour) can only

554   leave the soil column via the pore-space at the soil surface, or vegetated, where a considerable

555   proportion of the water being returned to the atmosphere consist of water taken up from the

556   deeper rooted parts of the soil profile.

557   *3.3.2. Outliers per scenario*

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

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

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

575 As some simulation runs did not converge (see discussion above), the comparison in terms of
576 total number of outliers is limited. Therefore, the total number of outliers was normalized to the
577 number of converged simulations for each scenario and PTF combination. Again, we present
578 the relative number of outliers for the homogeneous bare soil profile simulations in Fig. 5b as
579 an example (all others are shown in Annex Fig. 6 to 8). Here, the Weynants PTF shows the
580 largest percentage of outliers for the upper 0.95 and 0.85 percentile with 82 and 100 % outliers,
581 respectively. For the 'Toth class' PTF, we found 20 and 50 % outliers for the upper 0.95 and
582 0.85 percentile, respectively, indicating that also this PTF simulated larger fluxes with respect
583 to the ensemble. On the other hand, Rawls MvG shows the largest percentage of outliers at the
584 lower end (86 % for the 0.15 and 57 % for the 0.05 percentile) followed by Carsel&Parrish PTF
585 with 73 % for the 0.15 and 45 % for the 0.05 percentile. However, Rawls BC and Rawls class
586 also show substantial percentages of outliers for the 0.15 percentile. By comparing the relative
587 (converged only) and absolute (all runs) number of outliers, it can be seen that despite equal or
588 even lower or higher absolute number of outliers for different PTFs, the relative numbers differ
589 due to non-converged simulation runs for some PTFs. For instance, 'Toth class' for the 0.95
590 percentile showed 2 outliers yielding 20 % relative outliers as two simulations (silty clay and

591   silty clay loam) did not converge, whereby 1 outlier for Rosetta SSC yielded only 8 % relative

592   outliers as all simulations converged.

593   As there is no clear trend in the analysis of the absolute or relative outliers for the individual

594   scenarios (see Fig 5b and Annex Fig. 6b to 8b) which PTF generates most outliers, from here

595   on the outliers over all scenarios for all soil textural classes were calculated for converged

596   simulation runs only and expressed in relative terms. Figure 6a shows the outliers of the 90 %

597   tolerance interval (sum of upper and lower outliers), combined for all textural classes, for the

598   seven scenarios for the 13 PTFs for E_a and ET_a at t_{end}. In this figure, the PTFs of the two main

599   hydraulic formulations are clustered: those based on the Mualem van Genuchten (MvG) on the

600   left and those based on Brooks Corey (BC) formulation on the right. Furthermore, two lines are

601   added, dividing the results into three groups: i) those PTFs with relative number of outliers <

602   10 %, classified as 'robust', ii) those PTFs with 10 % ≥outliers ≤ 20 %, classified as

603   'intermediate robust', and iii) the PTFs with relative number of outliers >20 %, classified as

604   'non-robust'. It has to be noted that these thresholds (10 and 20 %) were chosen arbitrarily, but

605   may help to formulate the final recommendations for the choice of preferred PTF, to be used in

606   land surface models, for example.

607   This classification shows that the Rosetta SSC, Rosetta SSC+BD, Woesten, 'Toth continuous',

608   Rawls BC, Rawls class BC, Clapp&Hornberger, Cosby SC, and Cosby SSC PTFs are located

609   below the 10 % threshold for the 90 % tolerance interval, and can be therefore classified as

610   'robust' with respect to the ensemble behaviour (spread). Interestingly, all PTFs using BC

611   formulation show low relative numbers of outliers below 10%. Woesten PTF did not show any

612   outliers at all, indicating that this PTF is very robust with respect to the PTF ensemble used. On

613   the other hand, the 'Toth class' PTF was classified as intermediate robust, and three PTFs

614   (Carsel&Parrish, Rawls MvG, and Weynants) were classified as non-robust, whereby Rawls

615   MvG produced most outliers (32 %).

616   The results for the 70 % tolerance interval are shown in Fig. 6b and followed the same approach

617   as for the 90 % tolerance interval discussed above. Four PTFs are characterised as robust (Rawls

618   BC, Clapp&Hornberger, Cosby SC, and Cosby SSC). Again, all these four PTFs serve to

619   produce parameters for the Brooks Corey hydraulic formulation. The intermediate robust

620   grouping includes Rosetta SSC, Rosetta SSC+BC, Woesten, and 'Toth continuous', that

621   provide parameters for the Mualem van Genuchten formulation. Finally, Carsel&Parrish, Rawls

622   MvG, Weynants, 'Toth class', and Rawls BC class are those PTFs classified as non-robust.

623   There are two class, rather than continuous, PTFs here, indicating that continuous PTFs are

624   more likely to be robust. Also, the Weynants PTF was based on a relative small number of

625   samples, for Belgium only.

626   Based on the results presented above, it can be concluded that the use of different PTFs results

627   in different hydraulic properties that predict considerably different E_a or ET_a fluxes leading to

628   different soil water contents in the root zone but also to differences in deep percolation (or

629   ground water recharge). Furthermore, PTFs such as Carsel&Parrish, Rawls MvG and Weynants

630   can be identified as systematically less robust. In contrast, others, such as Woesten or all PTFs

631   using the Brooks Corey formulation (except Rawls BC class) seem to be robust with respect to

632   the ensemble of PTFs used in this study.

633   To facilitate the identification of outliers, all outliers per PTF, scenario, and textural class

634   combination were colour coded and plotted in Tab. 4. Again, Weynants overestimates E_a or ET_a

635   fluxes (brown colour for dryer soil conditions) for nearly all textural soil classes except for clay,

636   and silt. On the other hand, Rawls MvG shows underestimation (blue colour for wetter soil

637   conditions) for loam and silt loam over all three homogeneous soil scenarios and for silt and

638   sandy loam for two out of the three homogeneous soil scenarios. The Carsel&Parrish' PTF, on

639   the other hand, results in over- and underestimation, depending on soil class.

640   *3.3.3. Simulated spread with respect to scenario*

641   We raised the hypothesis that differences (variability) in simulated fluxes from using different

642   PTFs will be reduced with increasing model complexity. Increasing complexity was generated

643   by introducing vegetation (grass or wheat), soil layering, or the assumption of a fluctuating

644   ground water table, for the layered vegetated soil scenario only. As only the homogeneous

645   scenarios (bare, grass, and wheat) used all soil classes, we restrict the analysis on these three

646   scenarios.

647   For the analysis, again the simulated cumulative actual E_a or ET_a data at t_{end} was taken and the

648   model ensemble mean (MEM) for E_a ot ET_a at t_{end} over all PTFs was calculated for each

649   individual soil class and scenario. Based on the MEM value for E_a or ET_a, as well as the

650   individual E_a or ET_a value at t_{end} for each model run, the % difference from the MEM

651   ($100/MEM*E_a@t_{end}$ or $ET_a@t_{end}$) was calculated and visualized using boxplots in Fig. 7, where

652   the red line indicates the median, the box indicates the 0.25 and 0.75 percentiles, the whiskers

653   represent the most extreme data points not considered as outliers, and crosses represent the

654   outliers (value is more than 1.5 times the interquartile range). From the boxplots, two types of

655   information can be deduced: i), the variability of predicted E_a or ET_a over all PTFs for one soil

656   class / scenario and ii), the change in variability (spread) resulting from a change in scenario

657   complexity (bare, grass, or wheat vegetation).

658   In general, the largest variability in predicted E_a or ET_a was found for the bare soil conditions,

659   which is most pronounced for the loam, loamy sand, sand, sandy clay, clay loam, and sandy

660   loam class. Minor differences were found between bare and vegetated scenarios for the other

661   soil classes. The silty clay soil class for the grass scenario showed the smallest overall spread

662   between minimum and maximum predicted E_a (or ET_a) with a value of 7 % (min = 97 % and

663   max 104 %). On the other hand, the largest variability was found for the combination sandy

664   soil/bare soil scenario with 53 % (min = 72 % and max 125 %). All spreads, throughout the 13

665   PTFs, for different soil classes and scenarios are provided in the final column of Tab. 4. Overall,

666   bare scenarios show a mean spread of 30 %, whereby the grass and wheat vegetated scenarios

667   have only 23 % spread over all soil classes. A possible explanation for the reduced spread in

668   simulated E_a or ET_a with increasing model complexity (in this case vegetation) is that for the

669   vegetated profiles water is extracted from the rooted portion of the soil profile, whereas under

670   bare soil the water can only leave the soil profile at the soil surface. In the latter case, differences

671   in the soil hydraulic properties, especially in unsaturated hydraulic conductivity, which is

672   highly variable (on the order of magnitudes) between PTFs, close to the surface will impact the

673   E_a flux more substantially. As shown earlier, run off will occur in both scenarios (bare and

674   vegetated) and is even slightly larger for the vegetated scenario, and therefore, cannot explain

675   the reduced variability.

676   Next, for the layered soil scenarios (bare, grass, and wheat, without fluctuating groundwater

677   table) a clear reduction in the variability was observed, for both profiles (by sandy loam or silt

678   loam). The bare and the vegetated scenarios showed nearly the same spread (mean 13.4 % for

679   bare, 18.5 % for grass, and 15.5 % for the wheat). In general, the sandy loam overlaying silt

680   loam and loamy sand showed always higher variability compared to the silt loam overlaying

681   silty clay loam and silty clay, which is consistent to the finding that the sandy loam of the

682   homogeneous soil profile also showed higher variability compared to the homogeneous silt

683   loam scenarios.

684   Overall, the results indicate that adding vegetation reduces the variability in the simulated E_a

685   or ET_a flux, even if runoff occurs more frequently. This conclusion also holds for adding more

686   complexity in terms of soil layering, although the latter has to be regarded with some caution

687   due to the low number of soil combinations selected for these model runs. However, taking into

688   account that large portions of our global land surface is covered by vegetation, differences in

689   predicted fluxes, as a result of differences in PTFs used to generate the hydraulic parameters,

690   will most likely be smaller compared to an 'unvegetated world'.

691   In contrast, adding a fluctuating ground water table to the layered wheat scenario greatly

692   increased variability in ET_a flux, for both soil layering to 48 and 35 % for the sandy loam

693   overlaying silt loam and loamy sand, and silt loam overlaying silty clay loam and silty clay,

694   respectively.

695   *3.3.4. Differences in instantaneous fluxes*

696   Cumulative fluxes at t_{end} will only provide long-term systematic under or overestimation, but

697   will not provide information on how the instantaneous fluxes fluctuate compared to the MEM.

698   Therefore, the instantaneous fluxes were also analysed. The same analyses as conducted for the

699   cumulative fluxes were performed, i.e., calculation of the MEM and the 0.95 and 0.05

700   percentiles for time step i, whereby i runs from day 1 to 10988. Secondly, the total number as

701   well as the upper and lower percentile outliers were counted. As an example, the outliers of E_a

702   for the sandy loam of the homogeneous bare soil scenario were plotted in Fig. 8, for the different

703   PTFs. Carsel&Parrish PTF shows a substantial number of outliers for the lower 0.05 percentile

704   (N = 2020 or 18 % of all days), indicating that for these days less water will evaporate and

705   return to the atmosphere, which would have implications for the cloud forming processes of a

706   numerical weather prediction or climate model if a LSM using this PTF were to be embedded

707   within it. On the other hand, Weynants has 3053 outliers for the upper 0.85 percentile (28 %)

708   but also a smaller number of outliers for the lower 0.05 percentile (N = 309 or 3 %), leading to

709   larger E_a flux. A large number of 0.05 percentile outliers were also found for Rawls MvG (N =

710   2992 or 27 %), again combined with a lower number of upper 0.95 percentile outliers (N = 391

711   or 4 %). Cosby SC, Cosby SSC, and Rawls BC showed only low number of outliers (N < 10)

712   for the upper and lower percentiles. Even though the model runs for which the hydraulic

713   parameters were derived from Carsel&Parrish and Rawls MvG PTFs exhibit large numbers of

714   outliers, both are not flagged as 90 % tolerance interval outliers when the cumulative flux at

715   t_{end} was analysed. This means that the non-flagged instantaneous E_a fluxes compensate for the

716   lower fluxes determined as outliers in Fig 8, or that the outliers are close to the 0.15 percentile,

717   which is reflected by the fact that the total sum of underestimated flux (outlier flux – flux for

718   the lower 0.05 percentile for each outlier day) is low, amounting to 5.7 and 2.4 cm over the 30-

719   year period, respectively. Moreover, both PTFs show runoff exceeding a total of 1 cm in 60%

720   (Carsel&Parrish) and 36 % (Rawls MvG) of all converged simulations, respectively. For the

721   Rawls MvG the nine simulations with runoff even exceed the 100 cm threshold, with runoff

722   ranging between 388.6 to 859.2 cm. Looking at all textural classes (data not shown) for the

723   homogeneous bare soil scenario, 29 soil class / PTF combinations out of the total 151 do not

724   exhibit any outliers at all for the instantaneous E_a flux. These outliers are clustered in three soil

725   classes only (clay, silty clay, and silty clay loam). Interestingly, out of these 29 with zero

726   outliers in instantaneous flux, five are flagged as outliers for the cumulative flux at t_{end} (Rosetta

727   SSC clay, Cosby SSC clay, Weynants silty clay and silty clay loam, as well as Carsel&Parrish

728   silty clay loam), meaning that these PTFs over- or underestimate instantaneous E_a only very

729   modestly, yet consistently throughout the simulation period.

730   The percentage of all 90 percent tolerance outliers (sum of upper and lower outliers) summed

731   over all days for all three homogeneous soil scenarios (bare, grass, and wheat) for all soil classes

732   and PTFs are provided in Tab. 4. Over all soil classes and PTFs, the bare soil scenario has the

733   lowest total number of outliers ($N = 119930$ days or 6.5 % over all days and scenarios) followed

734   by the homogeneous wheat configuration ($N = 173961$ days or 10.1 %) and the homogeneous

735   grass scenario ($N = 178249$ days or 10.4 %). This finding is perhaps in contradiction to the

736   finding that the % spread in cumulative E_a or ET_a at t_{end} was larger for the bare soil scenario,

737   compared to the vegetated ones. Furthermore, for some texture classes the total number of

738   outliers increased remarkably when vegetation was implemented, such as for the clay class,

739   where the bare soil scenario has no outliers (0%), while the percentage of outliers increased to

740   14 % for the homogeneous grass and wheat scenario, respectively. This indicates that the

741   differences in available root zone water, affecting actual transpiration, are the main driver for

742   differences between PTFs, compared to fluxes over the soil surface E_a. On the other hand, only

743   the silty clay and the silty clay loam showed no outliers at all for the instantaneous flux for all

744   scenarios. Looking at all soil class/PTF/scenarios combinations, no clear trend in the total

745   number of outliers in instantaneous evapo(transpi)ration flux, and flagged outliers for the

746   cumulative E_a or ET_a flux at t_{end} can be observed. This leads to the conclusion that the outliers

747   in instantaneous flux alone do not necessarily sum up to a cumulative flux flagged as an outlier.

748   *3.4.Explaining variability and outliers by soil physical properties*

749   As has been shown, substantial variability exists in cumulative and instantaneous fluxes, and

750   some PTFs are found to be more robust than others. In this section, we discuss in more detail

751   the reasons for the differences between the predicted soil water fluxes, resulting from the use

752   of different PTFs, by analysing the estimated hydraulic parameters K_s, λ (MvG tortuosity

753   parameter) and the soil physical characteristics. In general, variability between estimated K_s for

754   the different PTFs is quite low (Fig. 9a), and values for Rawls MvG and BC only are

755   significantly lower than all other PTFs. These lower values may explain the poor numerical

756   convergence for these simulations, and the prevalence of lower E_a fluxes as well as a high
757   number of lower 0.05 percentile outliers at t_{end} as depicted in Tab. 4, especially for Rawls MvG.
758   Clapp&Hornberger K_s values are significantly higher than those estimated by the Weynants
759   PTF, 'Toth class', and Cosby SC and SSC, yet did not show any high outliers for E_a fluxes.
760   Interestingly, Cosby SC and SSC were developed based on the same water retention and K_s
761   data as Clapp&Hornberger, as both used data from Holtan et al. (1968), nevertheless estimated
762   K_s values are quite different. One reason might be that Clapp&Hornberger only used textural
763   classes, and averaged K_s for those classes, whereas Cosby SC and SSC is a continuous PTF.
764   Coming back to the outliers listed in Tab. 4, those runs based on Weynants PTF indicate larger
765   E_a fluxes at t_{end} and a large number of upper 0.95 percentile outliers, whereas their estimated
766   K_s is not significantly different from most other PTFs. Here, it has to be noted that Weynants
767   did not estimate K_s but rather estimated a near saturation hydraulic conductivity K_s* that is
768   mainly controlled by textural properties and which is lower that K_s. The results suggest that
769   variability in K_s alone cannot explain the flux differences simulated.

770

771   Looking at the λ value used in MvG formulation, two different classes of PTFs can be
772   distinguished, those setting λ to 0.5 as originally proposed by van Genuchten (1980)
773   (Carsel&Parrish, Rawls MvG, and 'Toth continuous') and those who fitted λ as an additional
774   free parameter (Rosetta SC and SSC, Woesten, Weynants, and 'Toth class'). The variability in
775   λ is plotted in Fig. 9b. It shows that the λ estimates of Weynants' PTFs are significantly lower
776   than those from the other four PTFs estimating λ, except for 'Toth class'. 'Toth class' λ values
777   are significantly lower than those calculated by Rosetta SC and SSC, and than those setting λ
778   to 0.5, whereas Woesten is significantly lower than Rosetta SC and SSC, and < 0.5. The more
779   negative λ values for Weynants appear strongly related to the larger number of upper 0.95
780   percentile outliers listed in Tab. 4, whereas the intermediately low λ values for 'Toth class' and
781   Woesten PTF do not explain the number of flagged outliers. In general, λ is significantly
782   correlated to the MvG parameter n for those PTFs setting $\lambda \neq 0.5$ ($R^2 =0.40$, $p = 0.05$, data not
783   shown) indicating a nonlinear behaviour which can be described as $n = 1.58\, e^{0.064\,\lambda}$ with an
784   R^2 of 0.51. Looking at the ranges of λ, there is a systematic difference between PTFs, with
785   largest λ values for Rosetta (-3.1>λ<0.62), followed by Woesten ((-4.46>λ<0.60), 'Toth class'
786   (-5.5>λ<0.73), and Weynants (-7.87>λ<1.92). Rosetta and Woesten are characterized by low
787   numbers of tolerance interval outliers, whereas 'Toth class' and Weynants are characterized by
788   large number of tolerance outliers, both in the upper end (upper 0.95 percentile outliers). As λ

789   is correlated to the *n* parameter, and *n* directly impacts the hydraulic properties and hence L_G,

790   L_C, τ_{FC}, and t_{FC}, and to a less extend S, the correlation between λ and these soil characteristics

791   was calculated. The results indicated (data not shown) that λ is not significantly correlated to

792   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 = $

793   0.05), whereas λ is not correlated to the flux E_a at t_{end}.

794

795   For the calculated soil characteristics L_G, Weynants shows large variability and high median

796   and significantly differs from Woesten, Rawls MvG, Rawls BC class, Rawls BC and Cosby SC

797   and SSC. In contrast, Rawls BC and BC class show low L_G, and Rawls BC class is significant

798   different from Rosetta SSC and Clapp&Hornberger (see Fig. 10a). Here, it has to be kept in

799   mind that L_G solely depends on the water retention characteristics and hence the *n* and α values

800   play a crucial role in the calculation. As *n* is positively correlated with λ, and Weynants shows

801   the smallest λ values, the significant difference, with regards to L_G, between Weynants and

802   most other PTFs seems logical. Large L_G values occur for very fine textures, which are

803   classically associated to low K values that limit water supply to the evaporating surface, which

804   is reflected by the higher number of upper 0.95 percentile outliers for Weynants, leading to a

805   drier soil profile. Lower E_a fluxes at t_{end}, and therefore, a wetter profile occurred frequently for

806   Carsel&Parrish and Rawls (MVG and BC), whereby all these PTFs are also located in the low

807   L_G range.

808   The calculation of L_C is based on knowledge of L_G and the actual hydraulic conductivity

809   distribution above the evaporation front. Therefore, K_s plays also an important role in the

810   calculation of L_C. The impact of K_s on L_C is clearly reflected in the high L_C values for Clapp &

811   Hornberger, which exhibit high K_s values across all soil classes compared to all other PTFs (see

812   Fig. 10b). At the other end of the spectrum, the impact of K_s on L_C is also apparent for Rawls

813   MvG and Rawls BC which do not indicate much spread and are characterized by low K_s and

814   hence low L_C. Surprisingly, Clapp&Hornberger are not classified as outliers when looking at

815   cumulative fluxes (see Tab. 4), whereas the low L_C for Rawls MvG corresponds to the number

816   of outliers detected. On the other hand, Weynants, which was characterized as the PTF with

817   most outliers at the upper 0.95 percentile, lies in the middle of the range of L_C values depicted

818   in Fig. 10b, indicating that L_C might not be a good indicator for flagged outliers. As stated in

819   Lehmann et al. (2008), L_C longer than 1 m are considered as unrealistic (evaporative extraction

820   of water by capillary flow across several meters is unlikely). Interestingly, only the

821   Clapp&Hornberger PTF show $L_C > 1$ m, while all other PTFs give realistic values.

822   The analysis of *MFP* shows a quite different picture (Fig. 10c). Here, the PTFs based on Brooks

823   Corey group together and exhibit a higher *MFP* compared to the MvG based PTFs. Testing on

824   significance showed that Rawls BC class, Clapp&Hornberger, and both Cosby PTFs are

825   significantly different from all others and that only Rawls BC is not significantly different from

826   those using MvG formulation, except for Rawls MvG. This is of interest, as Rawls MvG is only

827   a 'translation' of the Brooks Corey to van Genuchten parameters from Rawls BC according to

828   Morel-Seytoux (1986), while keeping K_s. As the Weynants PTF showed substantial outliers, as

829   listed in Tab. 4, one would also expect Weynants to be different with regards to *MFP* as the λ

830   value is much smaller compared to all other PTFs, while K_s does not differ (see Fig. 9a and 9b).

831   One reason for the fact that *MFP* for Weynants does not differ from the other PTFs. might be

832   its relatively low *n* value, as λ and *n* are positively correlated. The impact of λ as opposed to

833   the effect of MFP becomes clearer when we compare Weynants and Woesten, which show no

834   significant difference in *MFP*, yet larger K_s values for Woesten and lower λ for Weynants.

835   Overall, the *MFP* cannot explain the outliers detected and depicted in Tab. 4 as only Rawls

836   MvG is systematically different and exhibits large number of outliers, whereas Weynants *MFP*

837   are in the centre of the range of values found for the different PTFs. On the other hand, *MFP*

838   values for Clapp&Hornberger, as well as for both PTFs from Cosby, are significantly higher,

839   yet do not stand out in Tab. 4.

840   With regards to the sorptivity *S* (Fig. 11a), there is a large variability in *S* for

841   Clapp&Hornberger, which is significantly different from all other PTFs. Small variabilities in

842   *S*, however, are found for Woesten, Rawls MvG and BC, Weynants, 'Toth class' and both

843   Cosby PTF. In general, *S* is moderately correlated to L_C ($R^2 = 0.40$).

844   Rawls BC shows a high t_{grav}, which is significantly different from all other PTFs, except for

845   Rawls MvG. Both Cosby PTFs and both Rawls continuous functions (Rawls MvG and BC)

846   show relatively large variability (Fig. 11b). The higher t_{grav} for Rawls MvG fits with the larger

847   number of outliers listed in Tab. 4, whereas for BC this pattern is not clear, maybe due to the

848   lack of numerical convergence. In general, larger t_{grav} values are associated with more fine-

849   grained soils such as loam and clays (Alastal, 2012), whereas the low t_{grav} of Woesten

850   characterizes more coarse soils such as sands.

851   High τ_{FC} were calculated for Rawls MvG (Fig. 11c), whereby the large τ_{FC} is associated with

852   extremely low predicted K_s values. Extremely high values were found for Rawls MvG with τ_{FC}

853   exceeding 3 Mio. days, whereby Rawls MvG has K_s values of 0.01 and 0.004 cm d^{-1} for the

854   silty clay and clay class, respectively, and also did not converge. For the two soil classes, silt

855   and silt loam, where the model run did converge τ_{FC} is also extremely large (>44,000 days) and

856 for these soils again low K_s values of 0.2 and 0.3 cm d^{-1}, respectively, were estimated.

857 Additionally, these two model runs are also outliers at the lower 0.05 percentile.

858 Clapp&Hornberger PTF resulted in the smallest τ_{FC}, whereas the K_s predictions are in general

859 higher as for the other soils (see Fig. 9a) and none of the simulations were flagged as outliers.

860 On the other hand, all other PTFs have comparable τ_{FC} values, and the outliers detected in Tab.

861 4 seem not to be linked with τ_{FC}.

862 Finally, t_{FC} was analysed, which shows the same pattern as τ_{FC}, which is to be expected as t_{FC}

863 and τ_{FC} are linearly correlated as also shown by Assouline and Or (2014).

864 As these soil physical characteristics were calculated to help explain differences in simulated

865 E_a at t_{end}, all characteristics were correlated against E_a at t_{end} (see Fig. 12). Only $\log_{10}(L_G)$ shows

866 a moderate correlation to E_a at t_{end} ($R^2 = 0.52$, $p = 0.05$) and a weak correlation was found for

867 $\log_{10}(L_C)$, with $R^2 = 0.29$ ($p = 0.05$). As E_a, and also drainage D at t_{end}, will be biased if runoff

868 is generated (because less water will infiltrate into the soil profile and be available for

869 evaporation and drainage), E_a and drainage at t_{end} were normalized (E_{a_norm}, D_{norm}) by dividing

870 E_a or drainage at t_{end} by the difference of precipitation at t_{end} (2479.72 cm) and runoff at t_{end}. By

871 doing so, the correlation between λ and E_{a_norm} increased to $R^2 = 0.31$ ($p = 0.05$). For the derived

872 soil characteristics the correlation also increased (to $R^2 = 0.57$; $p = 0.05$) for $\log_{10}(L_G)$ but

873 decreased for $\log_{10}(L_C)$, to $R^2 = 0.10$. On the other hand, the correlation slightly increased for

874 t_{FC}, from $R^2 = 0.09$ to 0.22.

875 In a next step, a principal component analysis (PCA) using all converged model runs and soil

876 hydraulic parameters available for MvG and BC (θ_r, θ_s, K_s) as well as all soil characteristics

877 (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

878 performed on log transformed data (except θ_r, θ_s, MFP, E_{a_norm}, and D_{norm}) and the results are

879 plotted in Fig. 13. The first three components explain 76 % of the variability in the data and the

880 important loadings on PC 1 (42.9 % of variability) are t_{FC} (0.38), K_s (-0.35), and L_G (0.33). PC

881 2 (24.5 % of variability) includes the important loadings L_C (0.47), E_a at t_{end} (0.40), and S (0.31).

882 PC 3 explains only 8.6 % of the variability and t_{grav} (0.48) and θ_s (0.47) are the important

883 loadings. The PCA triplot shows scatter of the individual PTFs around the origin of the triplot

884 but also distinct PTF clusters, whereby Weynants (black circle) is oriented along the PC 1 in a

885 fairly small volume and is positively correlated to t_{FC} and τ_{FC} and negatively to D at t_{end} (as

886 drainage D at t_{end} is negative per definition). Rawls (MvG and BC) is oriented in the same

887 direction as Weynants but it exhibits larger scatter, whereas Clapp&Hornberger (red solid

888 markers) is oriented along PC 2 and correlates positively with K_s. K_s values reported by

889 Clapp&Hornberger are amongst the highest compared to all other PTFs as already discussed in

890 relation to Fig. 9a.

891 Out of these 13 PTFs, three (Clapp&Hornberger, Weynants, and Rawls) can be identified as

892 being distinctive from all others in the triplot as they do not cluster around the origin.

893 Furthermore, they do not only differ considerably in their estimated soil hydraulic parameters

894 (e.g., λ and n value for Weynants, and K_s for Rawls and Clapp&Hornberger) but also in the soil

895 characteristics derived from these parameters, whereby in all soil characteristics either the n

896 value (remember that n is correlated to λ) as well as K_s are directly or indirectly integrated. For

897 example, the low L_C values for Rawls PTFs indicate that the maximum extent of the flow region

898 sustaining evaporation is much smaller than for all other PTFs. This results in low E_a at t_{end}

899 compared to other PFTs and larger number of outliers as depicted in Tab. 4.

900 Finally, a multiple regression was performed to test whether E_a at t_{end} can be predicted by the

901 soil hydraulic parameters and/or characteristics, whereby only one of those parameters or

902 characteristics were used in turn, i.e. those that were available for MvG and BC. As per Fig. 13,

903 all entries were log transformed except for θ_r, θ_s, and MFP, and the best regression was selected

904 using bootstrapping. The best predictive model was found by E_a @ $tend$ = 1252.13 + 183.30

905 $\log_{10}(L_G)$ + 367.88 $\log_{10}(L_C)$ - 405.22 $\log_{10}(S)$ with an R^2 of 0.88 (see Fig. 14) pointing to the

906 fact that the soil characteristics L_G, L_C, and S describe well the physical behaviour of soils with

907 regards to actual evaporation. Using E_{a_norm} instead of E_a decreased the predictive power of the

908 multiple regression ($R^2 = 0.75$).

909

910 **4.   Summary and Conclusion**

911 In this study 13 pedotransfer functions (PTF) were used to populate the hydraulic parameters

912 required in the HYDRUS model that was then used to simulate the water fluxes for 12 USDA

913 soil classes, for different model scenarios that varied in complexity (homogeneous or layered

914 soil profile, with and without vegetation) over a period of 30 years. Plotting the hydraulic

915 functions (water retention and hydraulic conductivity curves) for all PTFs revealed large

916 differences, especially for the hydraulic conductivity curve, leading to the hypothesis that the

917 different PTFs will also show substantial differences in simulated fluxes.

918 It turned out that some PTFs generated parameters that rendered the HYDRUS model

919 numerically unstable, so that it failed to converge for certain soil class/configuration

920 combinations, especially those reported by Rawls and Brakensiek (1985) (Rawls MvG) and by

921 Rawls et al. (1982) (Rawls BC), which converged only in less than 44 % off all simulation runs.

922 Surprisingly, PTFs using the Brooks Corey (BC) formulation resulted in higher convergence

923   rates, compared to those based on Mualem van Genuchten, even though BC is in general

924   perceived to be less stable.

925   In a next step, differences in simulated actual evaporation E_a or evapotranspiration ET_a between

926   the model runs were analysed, as E_a and ET_a indirectly contain information on the net

927   infiltration, deep drainage (over long-term) and water stored in the root zone. Therefore, the

928   cumulative E_a or ET_a at the end of the simulation period (t_{end} = 10988 days) was selected and

929   the 90 and 70 % tolerance interval as well as the model ensemble mean were calculated. Fluxes

930   exceeding the tolerance limits were flagged and counted. The results indicate that some PTFs

931   (Rawls MvG, Weynants, and Carsel&Parrish) were classified as non-robust, as the fluxes

932   generated by the parameters derived from these PTFs exceeded a defined threshold of 20 % of

933   the 90 % tolerance interval outliers over all scenarios and soil classes. On the other hand, all

934   PTFs using the Brooks Corey formulation (Rawls BC, Rawls BC class, Clapp&Hornberger,

935   Cosby SC, and Cosby SSC) are classified as robust, as they generally result in a low percentage

936   of 90 % tolerance outliers. The PTF of Woesten performed best, and it showed no outliers at

937   all for the 90 % tolerance interval. A hypothesis raised at the beginning of the study was that

938   increasing model complexity will reduce the variability in predicted fluxes. Therefore, the

939   individual simulated E_a and ET_a fluxes at t_{end} were compared to the model ensemble mean

940   (MEM), and the relative spread of the individual simulations was calculated. The results show

941   that the bare soil scenarios exhibit the highest mean percentage spread (30 %), whereas the

942   grass and wheat vegetated scenarios had a reduced spread (23%), averaged over all soil classes.

943   The reduction in relative spread with the inclusion of vegetation can be explained by the fact

944   that for these runs the water leaving the soils can be extracted from the entire rooted soil profile

945   (after which it gets transpired via the vegetation), whereas under bare soil conditions it can only

946   leave the soil profile at the soil surface. In the latter case, differences in the soil hydraulic

947   properties close to the surface, especially in unsaturated hydraulic conductivity, which is highly

948   variable (in order of magnitudes) between PTFs, will impact the E_a or ET_a flux more

949   substantially.

950   Finally, the instantaneous E_a or ET_a fluxes over time were analysed, whereby again the 90 %

951   tolerance outliers were calculated and counted. The results indicate that some PTF/soil

952   class/model scenario combinations showed substantial outliers in the instantaneous fluxes, yet

953   were not flagged as outliers for the cumulative flux at t_{end}, indicating that the non-flagged

954   instantaneous fluxes compensate these outliers. On the other hand, other PTF/soil class/scenario

955   combinations showed no outliers for the instantaneous fluxes, but were flagged as outliers for

956 the cumulative case, indicating that even small over- or underestimations in instantaneous flux
957 can sum up to large errors in the long-run.

958 To explain differences in simulated E_a for the homogeneous bare soil scenario, different soil
959 characteristics were calculated, and a PCA was conducted using all simulated fluxes, soil
960 hydraulic parameters and soil characteristics available for both MvG and BC. The PCA revealed
961 three distinct PTFs clusters, namely Weynants, Rawls, and Clapp&Hornberger, whereby
962 Weynants and Rawls were also characterized by a large number of tolerance outliers. Weynants
963 correlates positively to gravity time of infiltration t_{FC} and τ_{FC} and negatively to drainage D at
964 t_{end}, whereby Clapp&Hornberger is oriented in the opposite direction and correlated with the
965 saturated conductivity K_s. For Rawls a reasonable correlation with t_{FC} and τ_{FC} is found, but due
966 to the large scatter for this PTF the interpretation is less clear.

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

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

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

985

989

990 **Data Availability Statement**

991 The HYDRUS-1D code is freely available online (https://www.pc-

992 progress.com/en/Default.aspx?h1d-downloads). The atmospheric data are also freely available

993 online (https://doi.org/10.7910/DVN/C0J5BB).

994

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.

1027   Donatelli, M., Wösten, J.H.M. & Belocchi, G. (2004). Evaluation of pedotransfer functions. In:
1028   Y. Pachepsky and W.J. Rawls (Eds.) Development of Pedotransfer functions in soil
1029   hydrology. Elsevier, Amsterdam. p. 357–362.

1030   Ek, M.B. & Holtslag, A.A.M. (2004). Influence of Soil Moisture on Boundary Layer Cloud
1031   Development. *Journal of Hydrometeorology*, 5, 86–99. https://doi.org/10.1175/1525-
1032   7541(2004)005<0086:IOSMOB>2.0.CO;2

1033   Espino, A., Mallants, D., Vanclooster, M. & Feyen, J. (1995). Cautionary notes on the use of
1034   pedotransfer functions for estimating soil hydraulic properties. *Agricultural Water*
1035   *Management Journal*, 29. 235-253.

1036   Feddes, R.A., Bresler, E., & Neuman, S.P. (1974). Field test of a modified numerical model for
1037   water uptake by root systems. *Water Resources Research*, 10. 1199-1206.

1038   Grant, C.D. & Groenevelt, P.H. (2015). Weighting the differential water capacity to account
1039   for declining hydraulic conductivity in a drying coarse-textured soil. *Soil Research*. 53,
1040   386–391. doi:10.1071/SR14258.

1041   Groh, J., Vanderborght, J., Pütz, T., Vogel, H.J., Gründling, R., Rupp, H., Rahmati, M.,
1042   Sommer, M., Vereecken, H., & Gerke, H.H. (2020). Responses of soil water storage and
1043   crop water use efficiency to changing climatic conditions: A lysimeter-based space-for-
1044   time approach. *Hydrology and Earth System Sciences*, 24 (3).
1045   https://doi.org/10.5194/hess-2019-411.

1046   Guber, A.K., Pachepsky, Y.A., van Genuchten, M.Th., Rawls, W.J., Simunek, J., Jacques, D.,
1047   Nicholson, T.J. & Cady, R.E. (2006). Field-Scale Water Flow Simulations Using
1048   Ensembles of Pedotransfer Functions for Soil Water Retention. *Vadose Zone Journal*,
1049   5(1). 234-247. doi: 10.2136/vzj2005.0111.

1050   Haghighi, E., Shahraeeni, E., Lehmann, P., & Or, D. (2013). Evaporation rates across a
1051   convective air boundary layer are dominated by diffusion. *Water Resources Research*,
1052   49(3). 1602-1610.

1053   Hoffmann, H., Zhao, G., Asseng, S., Bindi, M., Biernath, C., Constantin, J., Coucheney, E.
1054   Dechow, R., Doro, L., Eckersten, H., Gaiser, T., Grosz, B., Heinlein, F., Kassie, B.T.,
1055   Kersebaum, K.-C., Klein, C., Kuhnert, M., Lewan, E., Moriondo, M., Nendel, C.,
1056   Priesack, E., Raynal, H., Roggero, P.P., Rötter, R.P., Siebert, S., Specka, X. , Tao, F.
1057   Teixeira, E., Trombi, G., Wallach, D., Weihermüller, L., Yeluripati, J., & Ewert, F.
1058   (2016). Impact of Spatial Soil and Climate Input Data Aggregation on Regional Yield
1059   Simulations. *PLOS ONE.* doi:10.1371/journal.pone.0151782.

1060   Holtan, H.N., England, C.B., Lawless, G.P., & Schumaker, G.A. (1968). Moisture-tension data
1061   for selected soils on experimental water-sheds, Rep. ARS 41-144. Agric. Res. Serv.,
1062   Beltsville, Md. 609 pp.

1063   Iwema, J., Rosolem, R., Rahman, M., Blyth, E., & Wagener, T. (2017). Land surface model
1064   performance using cosmic-ray and point-scale soil moisture measurements for
1065   calibration, *Hydrology and Earth System Science*, 21. 2843–2861.
1066   https://doi.org/10.5194/hess-21-2843-2017.

1067   Jaynes, D.B. & Tyler, E.J. (1984). Using soil physical properties to estimate hydraulic
1068   conductivity. *Soil Science*, 138. 298-305.

1069   Kuhnert, M, Yeluripati, J., Smith, P., Hoffmann, H., van Oijen, M., Constantin, J., Coucheney,
1070   E., Dechow, R., Eckersten, H., Gaiser, T., Grosz, B., Haas, E., Kersebaum, K.-C., Kiese,
1071   R., Klatt, S., Lewan, E., Nendel, C., Raynal, H., Sosa, C., Specka, X., Teixeira, E.,
1072   Wang, E., Weihermüller, L., Zhao, G., Zhao, Z., Ogle, S., & Ewert, F. (2017). Impact
1073   analysis of climate data aggregation at different spatial scales on simulated net primary
1074   productivity for croplands. *European Journal of Agronomy*, 88. 41-52. doi: EURAGR-
1075   25545.

1076   KAK. (1994). Bodenkundliche Kartieranleitung. 4th edition. E. Schweizerbart'sche
1077   Verlagsbuchhandlung. Stuttgart.

1078   Koster, R.D. & Suarez, M.J. (2001). Soil Moisture Memory in Climate Models. *Journal of*
1079   *Hydrometeorolgy*, 2. 558–570. https://doi.org/10.1175/1525-
1080   7541(2001)002<0558:SMMICM>.

1081   Klute, A.J. (1952). Some theoretical aspects of the flow of water in unsaturated soils. *Soil*
1082   *Science of America Proceedings*, 16. 144–148.

1083   Latorre, B., Moret-Fernández, D., Lassabatere, L., Rahmati, M., López, M.V., Angulo-
1084   Jaramillo, R., Sorando, R., Comín, F. & Jiménez, J.J. (2018). Influence of the β
1085   parameter of the Haverkamp model on the transient soil water infiltration curve. *Journal*
1086   *of Hydrology*, 564. 222-229.

1087   Lehmann, P., Assouline, S. & Or, D. (2008). Characteristic lengths affecting evaporative drying
1088   of porous media. *Physical Review E*, 77(5). 056309. doi:10.1103/PhysRevE.77.056309.

1089   Lehmann, P., Merlin, O., Gentine, P., & Or, D. (2018). Soil texture effects on surface resistance
1090   to bare soil evaporation. *Geophysical Research Letters*, 45. 10,398–10,405.
1091   https://doi.org/10.1029/2018GL078803.

1092   Morel-Seytoux, H.J., Meyer, P.D., Nachabe, M., Touma, J., van Genuchten, M.T., & Lenhard,
1093   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.

1127   Rahmati M., Vanderborght, J., Šimunek, J., Vrugt, J.A., Moret-Fernández, D., Latorre, B.,
1128   Lassabatere, L., & Vereecken, H. (2020). Soil hydraulic properties estimation from one-
1129   dimensional infiltration experiments using characteristic time concept. *Vadose Zone*
1130   *Journal*, https://doi.org/10.1002/vzj2.20068 .

1131   Rawls, W.J. & Brakensiek, D.L. (1985). Prediction of soil water properties for hydrologic
1132   modeling. p 293-299. In: Jones, E.B. and T.J. Ward (eds.). Watershed Management in
1133   the Eighties. Proc. Irrig. Drain. Div., ASCE, Denver, CO. April 30 - May 1, 1985., 34:
1134   3293-3302.

1135   Rawls, W.J., Brakensiek, D.L., & Saxton, K.E. (1982). Estimating soil water properties.
1136   *Transactions ASAE*, 25(5). 1316-1320 and 1328.

1137   Schaap, M.G. (2004). Accuracy and uncertainty in PTF predictions. p. 33–43. In: Y. Pachepsky
1138   and W.J. Rawls (Eds.) Development of pedotransfer functions in soil hydrology.
1139   Elsevier, Amsterdam.

1140   Schaap, M.G., & Leij, F.J. (1998). Database-related accuracy and uncertainty of pedotransfer
1141   functions. *Soil Science*, 163. 765–779.

1142   Schaap, M.G., Leij, F.J., & van Genuchten, M.T. (2001). Rosetta: A computer program for
1143   estimating soil hydraulic parameters with hierarchical pedotransfer functions. *Journal*
1144   *of Hydrology*, 251. 163–176.

1145   Schaap, M.G. & van Genuchten, M.T. (2006). A Modified Mualem–van Genuchten
1146   Formulation for Improved Description of the Hydraulic Conductivity Near Saturation.
1147   *Vadose Zone Journal*, 5. 27-34. doi:10.2136/vzj2005.0005.

1148   Shao, Y. & Irannejad, P. (1999). On the Choice of Soil Hydraulic Models in Land-Surface
1149   Schemes. *Boundary-Layer Meteorology*, 90. 83–115.
1150   https://doi.org/10.1023/A:1001786023282

1151   Šimůnek, J. & van Genuchten, M.T. (2008). Modelling nonequilibrium flow and transport
1152   processes using HYDRUS. *Vadose Zone Journal*, 7. 782-797.

1153   Šimůnek, J., van Genuchten, M. T., & Šejna, M. (2008). Development and applications of the
1154   HYDRUS and STANMOD software packages and related codes. *Vadose Zone Journal*,
1155   7(2). 587-600.

1156   Seneviratne, S.I., Corti, T. Davin, E.L., Hirschi, M., Jaeger, E.B., Lehner, I., Orlowsky, B., &
1157   Teuling, A.J. (2010). Investigating soil moisture–climate interactions in a changing
1158   climate: a review. *Earth Science Reviews*, 99 (3–4). 125–161.

1159 Tietje, O. & Hennings, V. (1996). Accuracy of the saturated hydraulic conductivity prediction
1160   by pedo-transfer functions compared to the variability within FAO textural classes.
1161   *Geoderma*, 69 (1-2). 71-84.

1162 Tietje, O. & Tapkenhinrichs, M. (1993). Evaluation of pedo-transfer functions. *Soil Science*
1163   *Society of America Journal*, 57 (4). 1088-1095.

1164 Tóth, B., Weynants, M., Nemes, A., Makó, A., Bilas, G., & Tóth, G. (2015). New generation
1165   of hydraulic pedotransfer functions for Europe. *European Journal of Soil Science*, 66.
1166   226–238. doi: 10.1111/ejss.12192.

1167 van den Hurk, B., Ettema, J., & Viterbo, P. (2008). Analysis of Soil Moisture Changes in
1168   Europe during a Single Growing Season in a New ECMWF Soil Moisture Assimilation
1169   System.   *Journal   of   Hydrometeorolgy*,   9.   116–131.
1170   https://doi.org/10.1175/2007JHM848.1.

1171 van Genuchten, M.Th. (1980). A closed form equation for predicting the hydraulic conductivity
1172   of unsaturated soils. *Soil Science Society of America Journal*, 44. 892 898.

1173 van Genuchten, M.T. & Pachepsky, Y. (2011). Pedotransfer functions. In: J. Glinski, J.
1174   Horabik, and J. Lipiec (Eds.). Encyclopedia of Agrophysics. Springer: 556-561. doi:
1175   10.1007/978-90-481-3585-1_109.

1176 van Looy K., Bouma, J., Herbst, M., Koestel, J., Minasny, B., Mishra, U., Montzka, C., Nemes,
1177   A., Pachepsky, Y., Padarian, J., Schaap, M., Tóth, B., Verhoef, A., Vanderborght, J.,
1178   van der Ploeg, M., Weihermüller, L., Zacharias, S., Zhang, Y. & Vereecken, H. (2017).
1179   Pedotransfer functions in Earth system science: challenges and perspectives. *Reviews of*
1180   *Geophysics*. DOI:10.1002/2017RG000581.

1181 Vereecken, H., Diels, J. Vanorshoven, J., Feyen, J., & Bouma, J. (1992). Functional evaluation
1182   of pedotransfer functions for the estimation of soil hydraulic properties. *Soil Science*
1183   *Society of America Journal*, 56. 1371–1378.

1184 Vereecken, H., Huisman, J.A., Bogena, H., Vanderborght, J., Vrugt, J.A., & Hopmans, J.W.
1185   (2008). On the value of soil moisture measurements in vadose zone hydrology: a review.
1186   *Water Resources Research*, 44. W00D06, doi:10.1029/2008WR006829. 21.

1187 Vereecken, H., Huisman, J.A., Pachepsky, Y., Montzka, C., van der Kruk, J., Bogena, H.,
1188   Weihermüller, L., Herbst, M., Martinez, G., & Vanderborght, J. (2014). On the spatio-
1189   temporal dynamics of soil moisture at the field scale. *Journal of Hydrology*, 516, 76–
1190   96.

1191 Vereecken, H., Maes, J., & Feyen, J. (1990). Estimating unsaturated hydraulic conductivity
1192   from easily measured soil properties. *Soil Science*, 149. 1–12.

1193 Vereecken, H., Maes, J., Feyen, J., & Darius, P. (1989). Estimating the soil moisture retention
1194   characteristic from texture, bulk density, and carbon content. *Soil Science*, 148. 389–
1195   403.

1196 Vereecken, H., Weynants, M., Javaux, M., Pachepsky, Y., Schaap, M.G., & van Genuchten,
1197   M.T. (2010). Using pedotransfer functions to estimate the Van Genuchten-Mualem soil
1198   hydraulic properties- a review. *Vadose Zone Journal*, 9. 795–820.
1199   doi:10.2136/vzj2010.0045

1200 Vereecken, H., Weihermüller, L., Assouline, S., Šimůnek, J., Verhoef, A., Herbst, M., Archer,
1201   N., Mohanty, B., Montzka, C., Vanderborght, J., Balsamo, G., Bechtold, M., Boone, A.,
1202   Chadburn, S., Cuntz, M., Decharme, B., Ducharne, A., Ek, M., Garrigues, S., Goergen,
1203   K., Ingwersen, J., Kollet, S., Lawrence, D.M., Li, Q., Or, D., Swenson, S., de Vrese, P.,
1204   Walko, R., Wu, Y., & Xue, Y. (2019). Infiltration from the Pedon to Global Grid Scales:
1205   An Overview and Outlook for Land Surface Modeling. *Vadose Zone Journal*, 18(1).
1206   doi: 10.2136/vzj2018.10.0191.

1207 Vogel, T., van Genuchten, M.Th., & Cislerova, M. (2001). Effect of the shape of the soil
1208   hydraulic functions near saturation on variably-saturated flow predictions. *Advances in*
1209   *Water Resources*, 24 (2). 133-144 https://doi.org/10.1016/S0309-1708(00)00037-3.

1210 Wagner, B., Tarnawski, V.R., Hennings, V., Müller, U., Wessolek, G., & Plagge, R. (2001).
1211   Evaluation of pedo-transfer functions for unsaturated soil hydraulic conductivity using
1212   an independent data set. *Geoderma*, 102. 275–297.

1213 Weihermüller, L., Herbst, M., Javaux, M., & Weynants, M. (2017). Erratum to "Revisiting
1214   Vereecken Pedotransfer Functions: Introducing a Closed-Form Hydraulic Model.
1215   *Vadose Zone Journal*, doi:10.2136/vzj2008.0062er

1216 Wesseling, J.G., Eibers, J.A., Kabat, P., & van den Broek, B.J. (1991). SWATRE, Instructions
1217   for input. Internal note, Win and Staring Centre, Wageningen.

1218 Weynants, M., Vereecken, H., & Javaux, M. (2009). Revisiting Vereecken Pedotransfer
1219   Functions: Introducing a Closed-Form Hydraulic Model. *Vadose Zone Journal*, 8 (1).
1220   86-95. doi: 10.2136/vzj2008.0062.

1221 Wösten J.H.M., Lilly, A., Nemes, A., & Le Bas, C. (1999). Development and use of a database
1222   of hydraulic properties of European soils. *Geoderma*, 90.169-185.

1223 Wösten, J.H.M., Pachepsky, Y.A., & Rawls, W.J. (2001). Pedotransfer functions: Bridging the
1224   gap between available basic soil data and missing soil hydraulic characteristics. *Journal*
1225   *of Hydrology*, 251. 123–150.

1226   Yakirevich, A., Pachepsky, Y.A., Gish, T.J., Guber, A.K., Kuznetsov, M.Y., Cady, R.E., &

1227   Nicholson, T. (2013). Augmentation of groundwater monitoring networks using

1228   information theory and ensemble modeling with pedotransfer functions. *Journal of*

1229   *Hydrology*, 501. 13-24. doi: 10.1016/j.jhydrol.2013.07.032.

1230

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].

1276
1277

1278 **Table Captions**

1279 **Table 1:** Overview of used pedotransfer functions (PTFs) with input parameters. MvG =
1280 Mualem- van Genuchten type (Eq. 4-5), BC = Brooks-Corey type (Eq. 1-3 Class =
1281 class transfer functions of USDA soil classes. Continuous = based on discrete soil
1282 information (see PTF Inputs), BD = bulk density (g cm^{-3}), pH (here set to pH = 6.5),
1283 and cation exchange capacity CEC (here set to CEC = 0.5*clay+0.005*silt according
1284 to KAK (1994)).
1285

1286 **Table 2:** Feddes root water uptake parameters for pasture (here used for the grass scenario) and
1287 wheat vegetation according to Wesseling (1991).

1288 **Table 3:** Overview of converged simulations for the different PTFs listed in Tab.1. Last two
1289 columns indicate the use of air entrance value of -2 cm for the Mualem- van
1290 Genuchten type function if simulation did not converge using predicted PTF
1291 parameters.

1292 **Table 4:** Overview of 90% tolerance interval outliers per textural class and scenario
1293 (homogeneous bare, grass, or wheat) for E_a/ET_a at t_{end} = 10988 days and the spread
1294 over all PTFs from the model ensemble mean (MEM) (colour coded in blue and
1295 brown). Numbers for the individual pedotransfer / soil class / scenario combinations
1296 depict the % of total 90 % tolerance interval outliers for the instantaneous E_a/ET_a
1297 flux. NaN are non-converged simulations. % spread is the spread in % between
1298 minimum and maximum cumulative E_a/ET_a at t_{end} over one soil class / scenario
1299 combination.
1300

1301

1302

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	θ$_{(h)}$	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							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 (θ$_{(h)}$) 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%
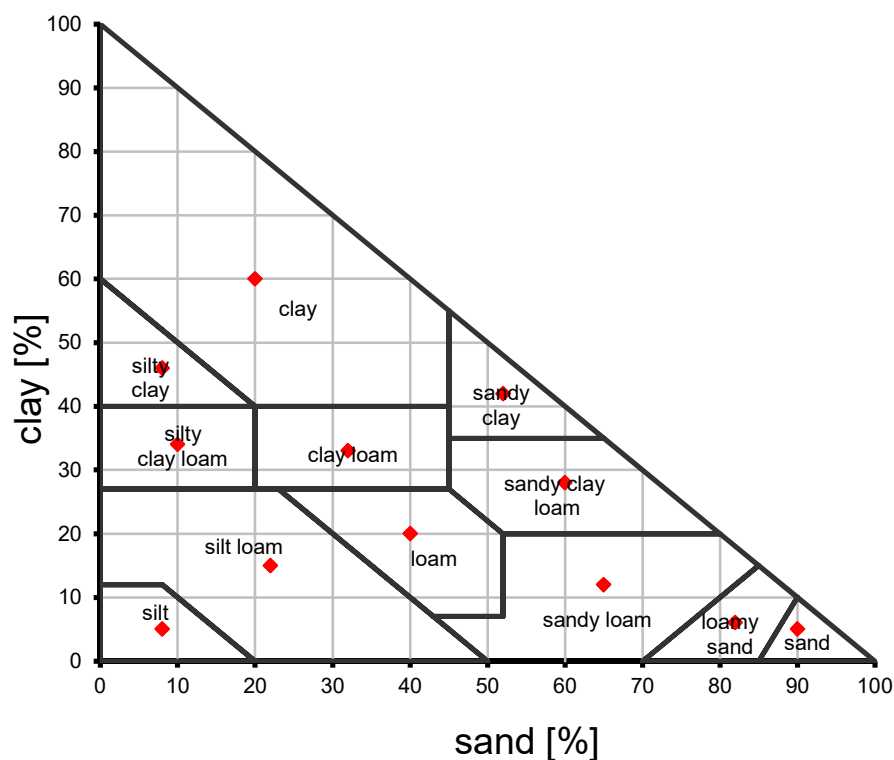■ class 10-20 %
■ class 20-30 %
■ class > 30 %

1335
1336
1337 **Figure 1:** Overview of the used soil textures for the 12 USDA soil classes. Red points indicate
1338   mean texture for each soil class.

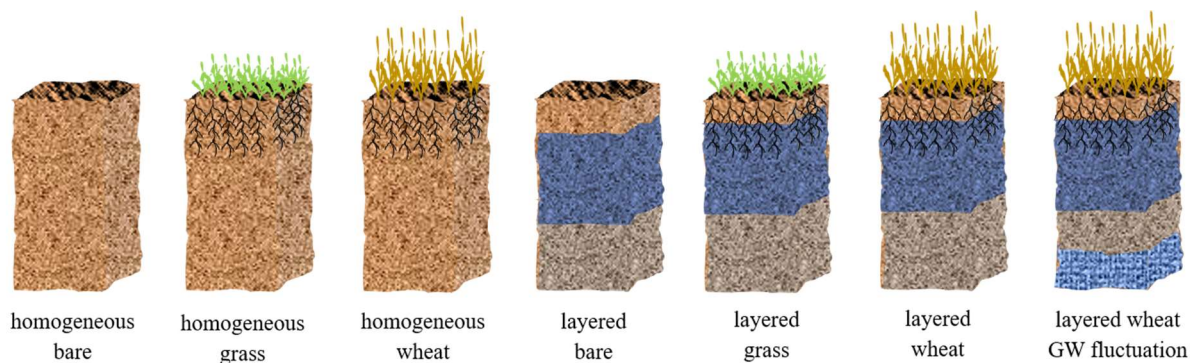1339

1340

1341

1342 **Figure 2:** Schematic sketch of the seven different scenarios used for the modelling study with
1343   increasing model complexity from left to right. The first three scenarios were
1344   computed for each soil textural class. The four layered scenarios were run for two
1345   different types of layering.

1346

1347

1348

1349

1350

1351

1352

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.

1353
1354
1355

1356

1357

1358

1359

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

1364

1365

1366

1367

1368

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.

1374

1375

1376

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.

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.

1385

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.

a)

b)

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.
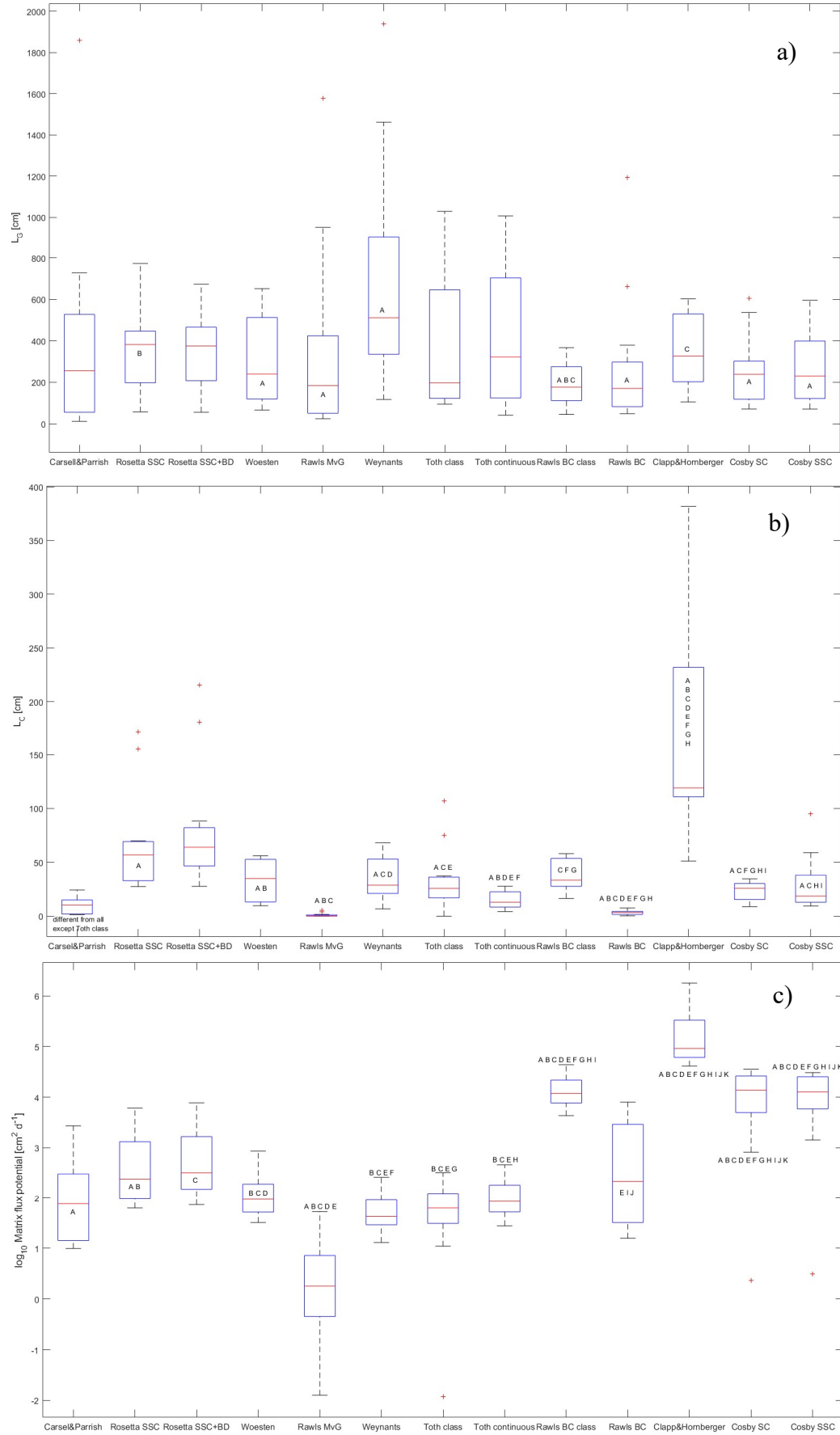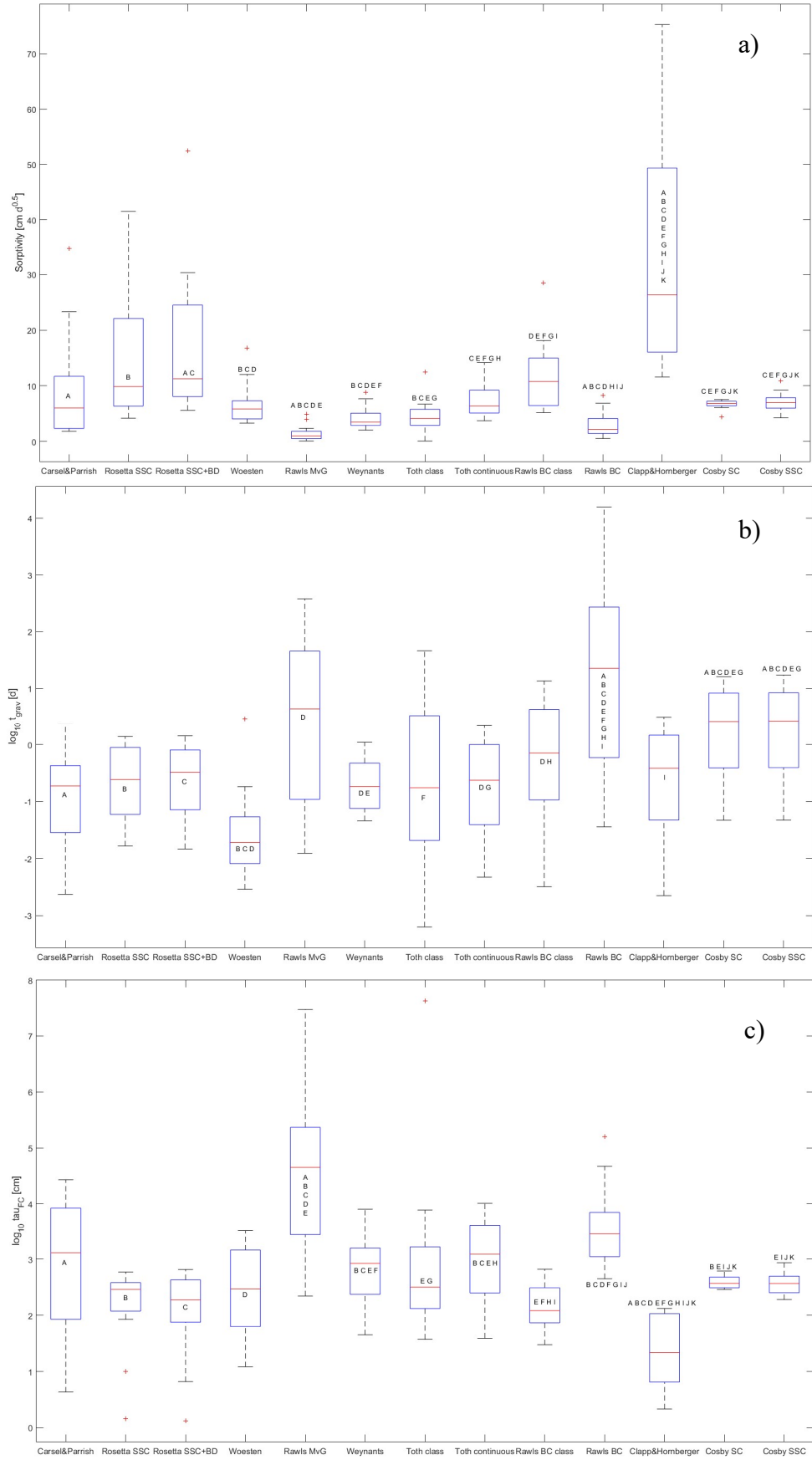
1390

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.
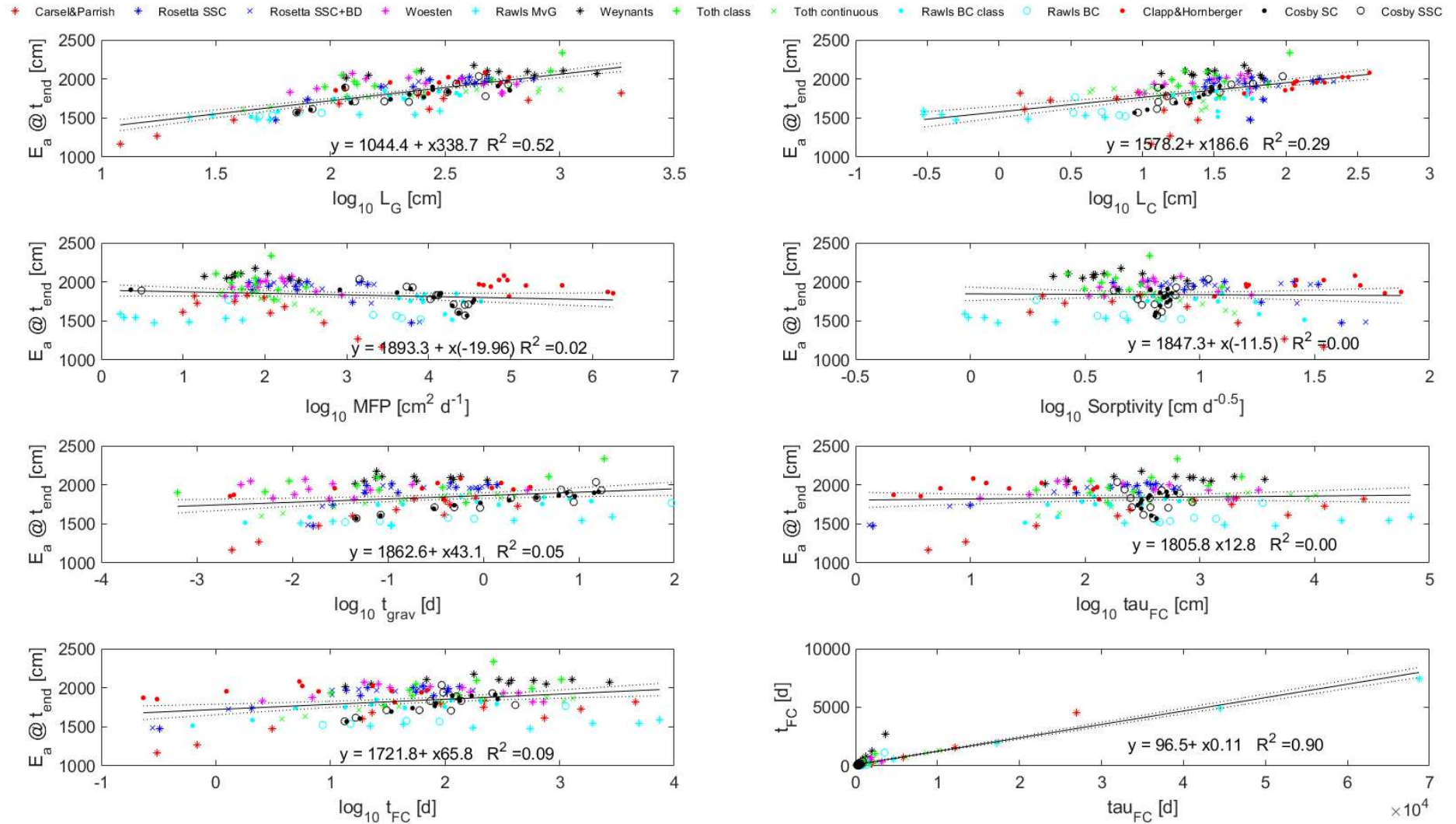
Figure 12: Scatterplots of the different soil characteristics gravitational length L_G, characteristic length of evaporation L_C, matrix flux potential *MFP*, sorptivity S, characteristic time t_{grav}, elapsed time for the attainment of field capacity t_{FC}, characteristic time for the attainment of field capacity τ_{FC}, versus E_a at t_{end} for the homogeneous bare soil scenario as well as τ_{FC} versus t_{FC}.
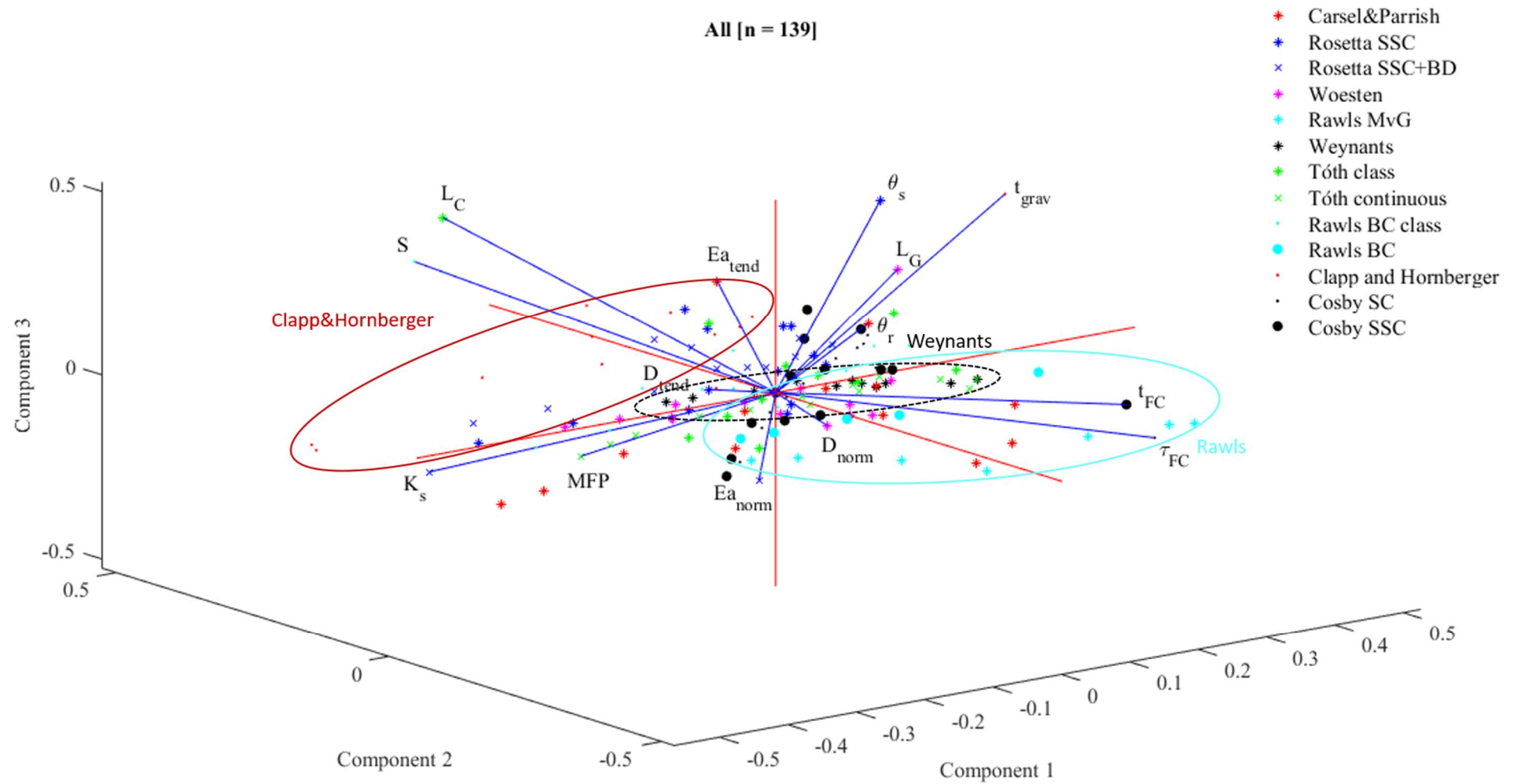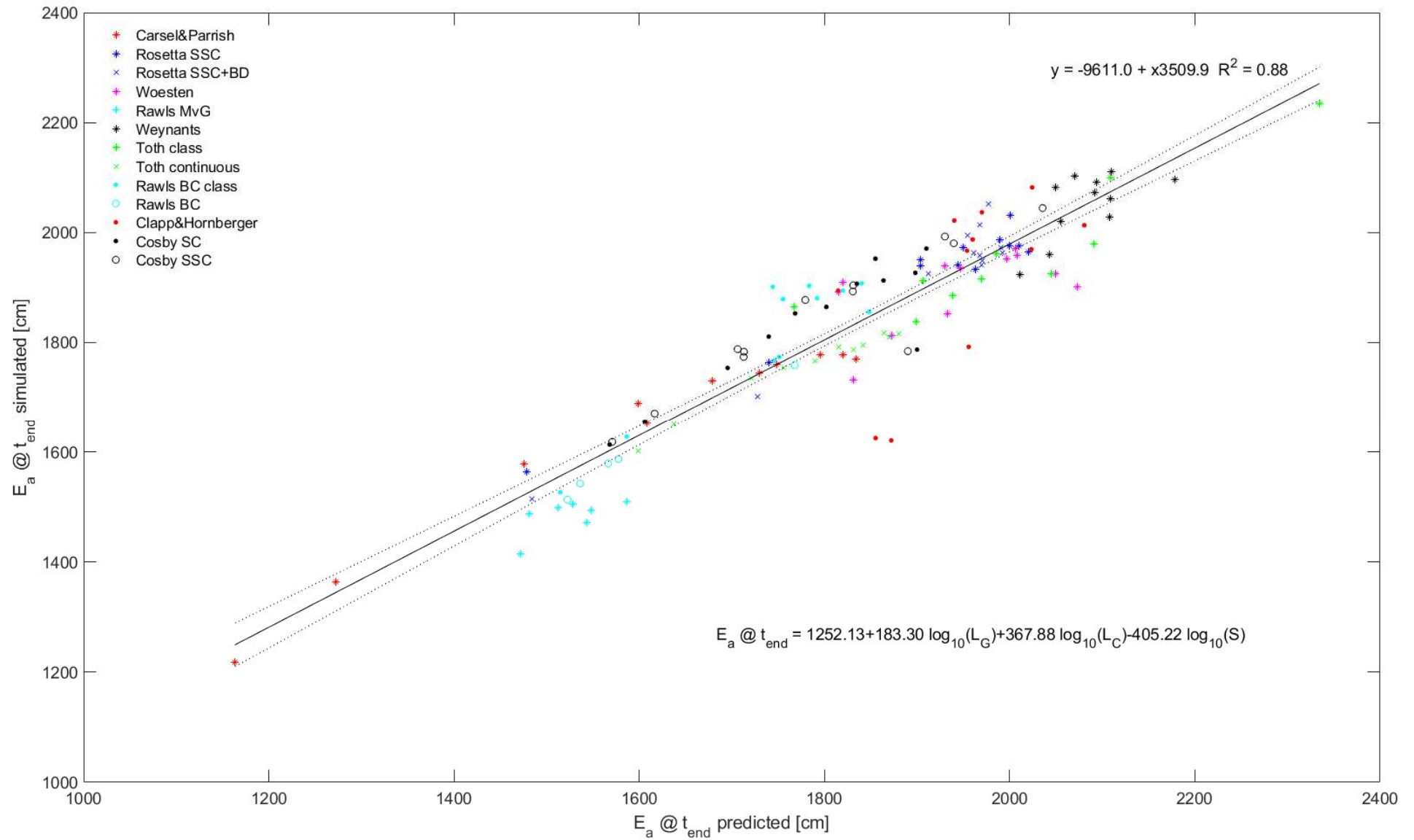
1400

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.
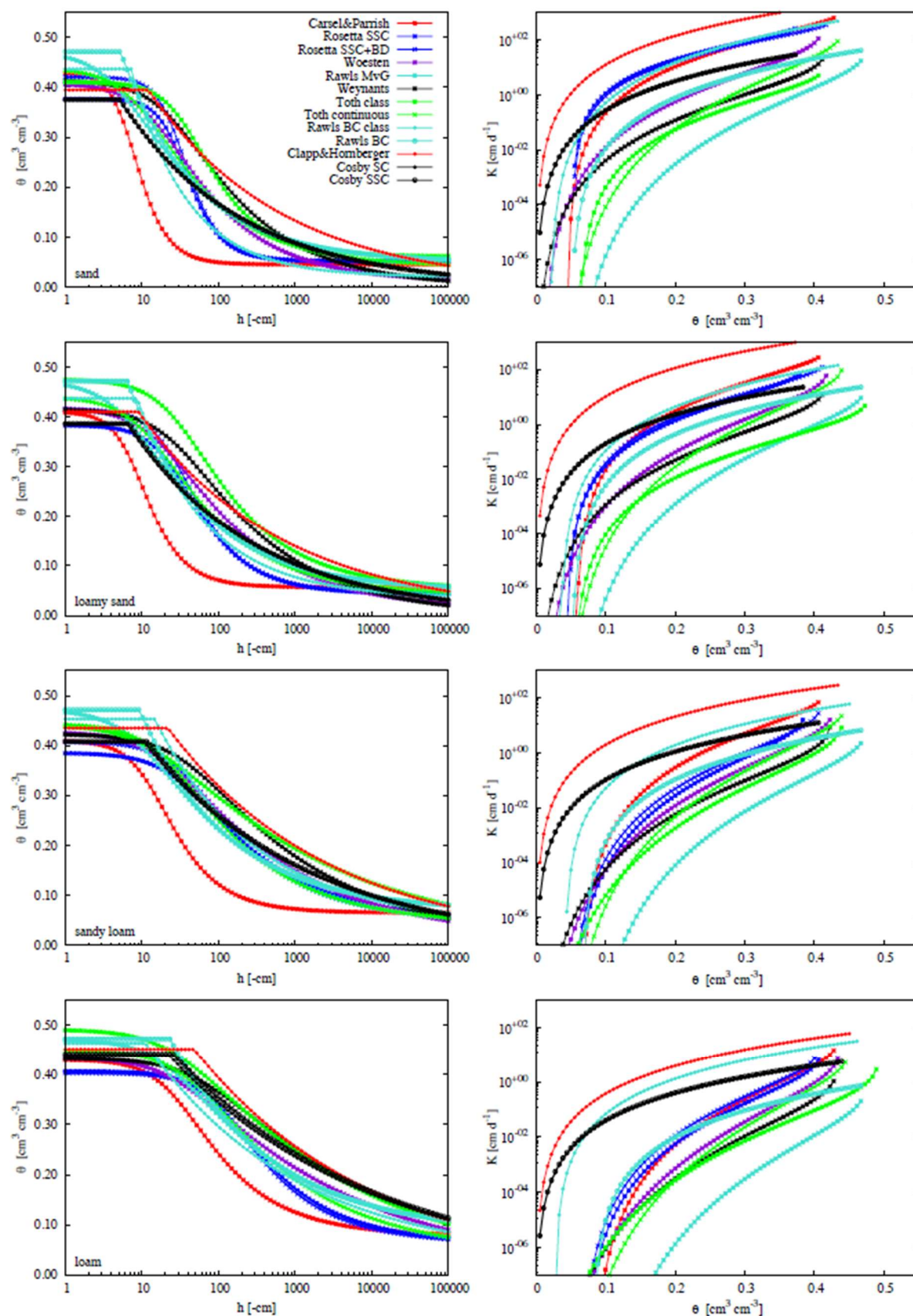
Class Clay	θ_r [cm³ cm⁻³]	θ_s [cm³ cm⁻³]	α [cm⁻¹]	n [-]	K_s [cm d⁻¹]	λ [-]
Carsel&Parrish (Carsel and Parrish, 1988)	0.068	0.380	0.0080	1.090	4.8	0.50
Rosetta SSC (Schaap et al., 2001)	0.097	0.485	0.0210	1.206	17.7	-2.59
Rosetta SSC+BD (Schaap et al., 2001)	0.010	0.475	0.0191	1.252	11.4	-1.83
Woesten (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

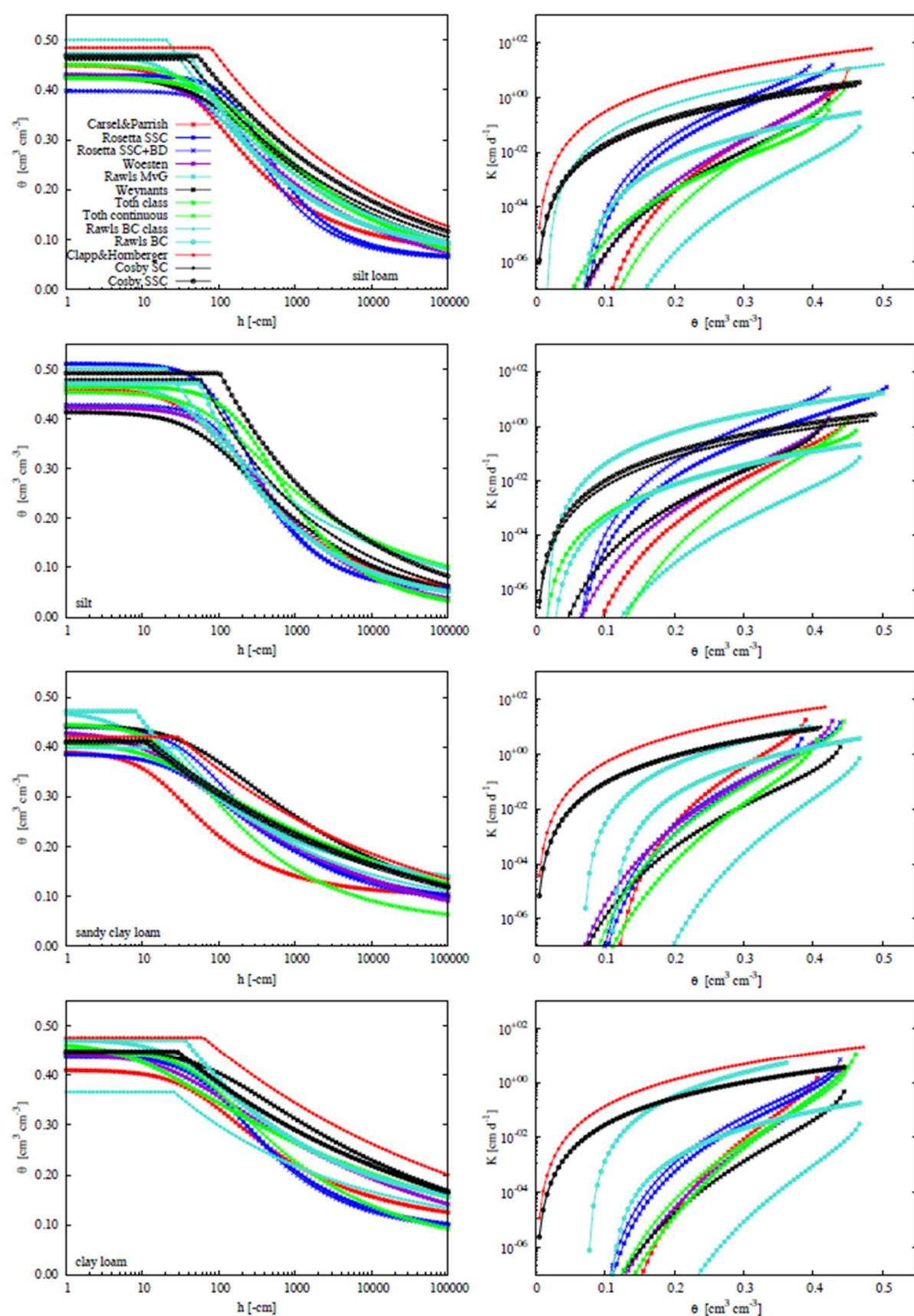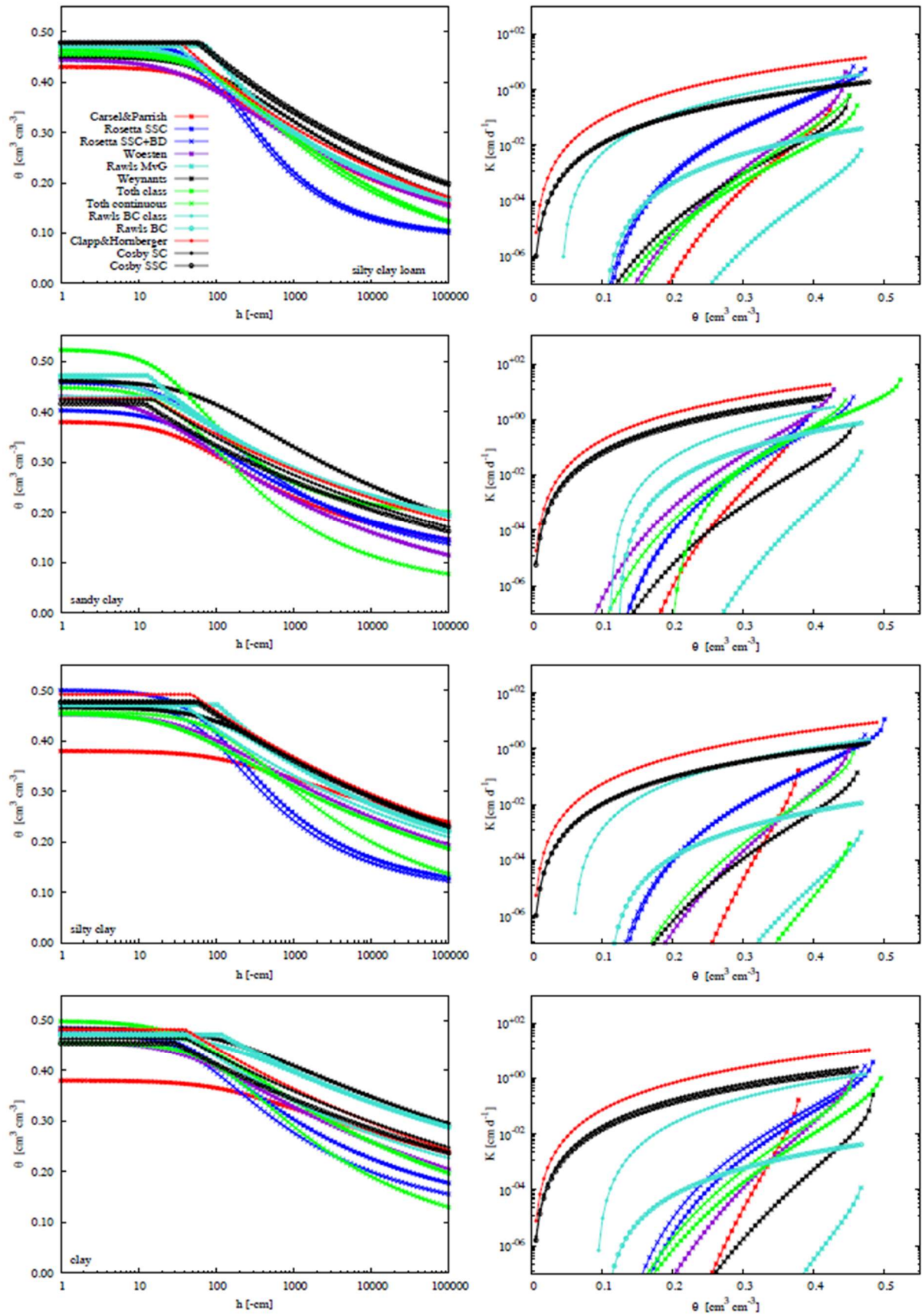| Class Sandy clay loam	θ_r [cm³ cm⁻³]	θ_s [cm³ cm⁻³]	α [cm⁻¹]	n [-]	K_s [cm d⁻¹]	λ [-]
Carsel&Parrish (Carsel and Parrish, 1988)	0.100	0.390	0.0590	1.480	31.4	0.50
Rosetta SSC (Schaap et al., 2001)	0.070	0.386	0.0269	1.290	12.5	-1.79
Rosetta SSC+BD (Schaap et al., 2001)	0.075	0.441	0.0216	1.370	32.0	-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*.

1410

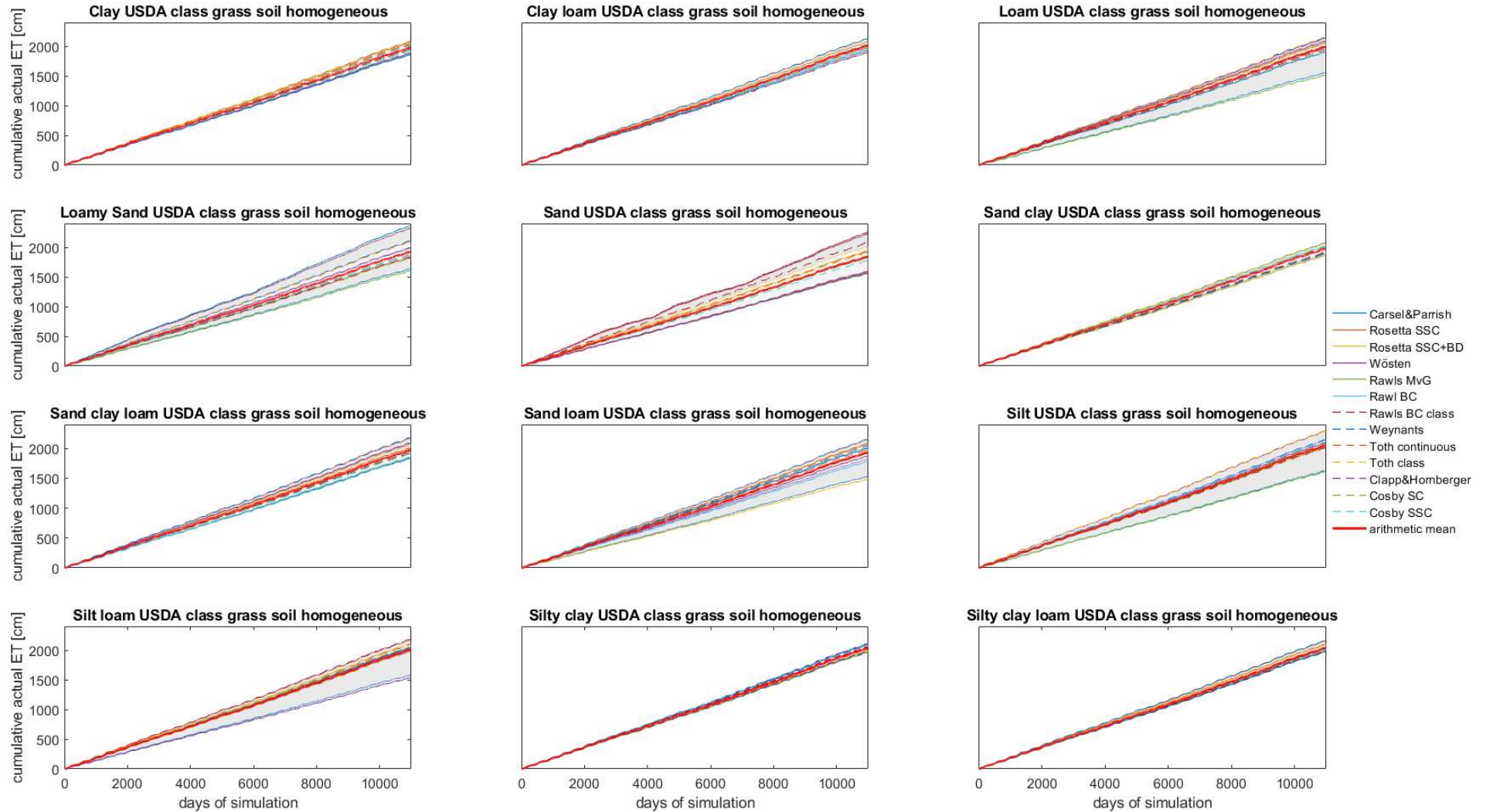| Class Clay	θ_r [cm³ cm⁻³]	θ_s [cm³ cm⁻³]	α [cm⁻¹]	b [-]	K_s [cm d⁻¹]	Class Sandy clay loam	θ_r [cm³ cm⁻³]	θ_s [cm³ cm⁻³]	α [cm⁻¹]	b [-]	K_s [cm d⁻¹]
Rawls BC class (Rawls et al., 1982)	0.090	0.475	0.0268	7,634	1.4	Rawls BC class (Rawls et al., 1982)	0.068	0.398	0.0356	4.000	10.3
Rawls BC (Rawls and Brakensiek, 1985)	0.101	0.472	0.0086	9.883	0.004	Rawls BC (Rawls and Brakensiek, 1985)	0.107	0.472	0.1241	3.893	4.0
Clapp&Hornberger (Clapp and Hornberger, 1978)	0,000	0.482	0.0247	11.400	11.1	Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.420	0.0334	7.120	54.4
Cosby SC (Cosby et al., 1984)	0,000	0.464	0.0241	12.450	2.5	Cosby SC (Cosby et al., 1984)	0.000	0.413	0.0805	7.362	10.2
Cosby SSC (Cosby et al., 1984)	0,000	0.454	0.0334	12.460	1.8	Cosby SSC (Cosby et al., 1984)	0.000	0.409	0.0900	7.316	9.0
Class Clay loam						**Class Sandy loam**					
Rawls BC class (Rawls et al., 1982)	0.075	0.366	0.0386	5.155	5.5	Rawls BC class (Rawls et al., 1982)	0.041	0.453	0.0682	3.106	62.2
Rawls BC (Rawls and Brakensiek, 1985)	0.105	0.472	0.0276	4.110	0.2	Rawls BC (Rawls and Brakensiek, 1985)	0.067	0.472	0.1087	2.689	7.1
Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.476	0.0159	8.520	21.2	Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.435	0.0459	4.900	299.5
Cosby SC (Cosby et al., 1984)	0.000	0.449	0.0346	8.157	3.8	Cosby SC (Cosby et al., 1984)	0.000	0.407	0.0936	4.818	12.2
Cosby SSC (Cosby et al., 1984)	0.000	0.447	0.0350	8.185	3.7	Cosby SSC (Cosby et al., 1984)	0.000	0.408	0.0856	4.789	13.1
Class Loam						**Class Silt**					
Rawls BC class (Rawls et al., 1982)	0.027	0.463	0.0897	4.545	31.7	Rawls BC class (Rawls et al., 1982)	0.015	0.501	0.0482	4.739	16.3
Rawls BC (Rawls and Brakensiek, 1985)	0.082	0.472	0.0428	3.116	0.8	Rawls BC (Rawls and Brakensiek, 1985)	0.020	0.472	0.0178	2.797	0.2
Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.451	0.0209	5.390	60.1	Clapp&Hornberger (Clapp and Hornberger, 1978)	NR	NR	NR	NR	NR
Cosby SC (Cosby et al., 1984)	0.000	0.439	0.0441	6.090	5.1	Cosby SC (Cosby et al., 1984)	0.000	0.479	0.0168	3.705	1.6
Cosby SSC (Cosby et al., 1984)	0.000	0.441	0.0387	6.120	5.6	Cosby SSC (Cosby et al., 1984)	0.000	0.492	0.0097	3.861	2.8
Class Loamy sand						**Class Silt loam**					
Rawls BC class (Rawls et al., 1982)	0.035	0.437	0.1150	2.110	146.6	Rawls BC class (Rawls et al., 1982)	0.015	0.501	0.0482	4.739	16.3
Rawls BC (Rawls and Brakensiek, 1985)	0.051	0.472	0.1548	2.343	23.9	Rawls BC (Rawls and Brakensiek, 1985)	0.062	0.472	0.0236	2.939	0.3
Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.410	0.1111	4.380	1350.7	Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.485	0.0127	5.300	62.2
Cosby SC (Cosby et al., 1984)	0.000	0.386	0.1564	3.864	22.2	Cosby SC (Cosby et al., 1984)	0.000	0.461	0.0256	5.295	2.7
Cosby SSC (Cosby et al., 1984)	0.000	0.386	0.1457	3.796	23.5	Cosby SSC (Cosby et al., 1984)	0.000	0.468	0.0187	5.389	3.6
Class Sand						**Class Silty clay**					
Rawls BC class (Rawls et al., 1982)	0.020	0.437	0.1380	1.689	504.0	Rawls BC class (Rawls et al., 1982)	0.056	0.479	0.0292	7.874	2.2
Rawls BC (Rawls and Brakensiek, 1985)	0.050	0.472	0.1905	2.254	43.7	Rawls BC (Rawls and Brakensiek, 1985)	0.105	0.472	0.0096	5.914	0.01
Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.395	0.0826	4.050	1520.6	Clapp&Hornberger (Clapp and Hornberger, 1978)	0,000	0.492	0.0204	10.400	9.0
Cosby SC (Cosby et al., 1984)	0.000	0.376	0.1991	3.705	29.4	Cosby SC (Cosby et al., 1984)	0,000	0.479	0.0168	1.224	1.6
Cosby SSC (Cosby et al., 1984)	0.000	0.376	0.1921	3.615	30.0	Cosby SSC (Cosby et al., 1984)	0,000	0.477	0.0176	1.298	1.5
Class Sandy clay						**Class Silty clay loam**					
Rawls BC class (Rawls et al., 1982)	0.109	0.430	0.0343	5.952	2.9	Rawls BC class (Rawls et al., 1982)	0.040	0.471	0.0307	6.623	3.6
Rawls BC (Rawls and Brakensiek, 1985)	0.120	0.472	0.0750	5.814	0.8	Rawls BC (Rawls and Brakensiek, 1985)	0.099	0.472	0.0130	4,226	0.04
Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.426	0.0654	10.400	18.7	Clapp&Hornberger (Clapp and Hornberger, 1978)	0.000	0.477	0.0281	7.750	14.7
Cosby SC (Cosby et al., 1984)	0.000	0.423	0.0633	9.588	7.7	Cosby SC (Cosby et al., 1984)	0.000	0.476	0.0178	8.316	1.8
Cosby SSC (Cosby et al., 1984)	0.000	0.416	0.0825	9.538	5.8	Cosby SSC (Cosby et al., 1984)	0.000	0.478	0.0159	8.408	1.9

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.
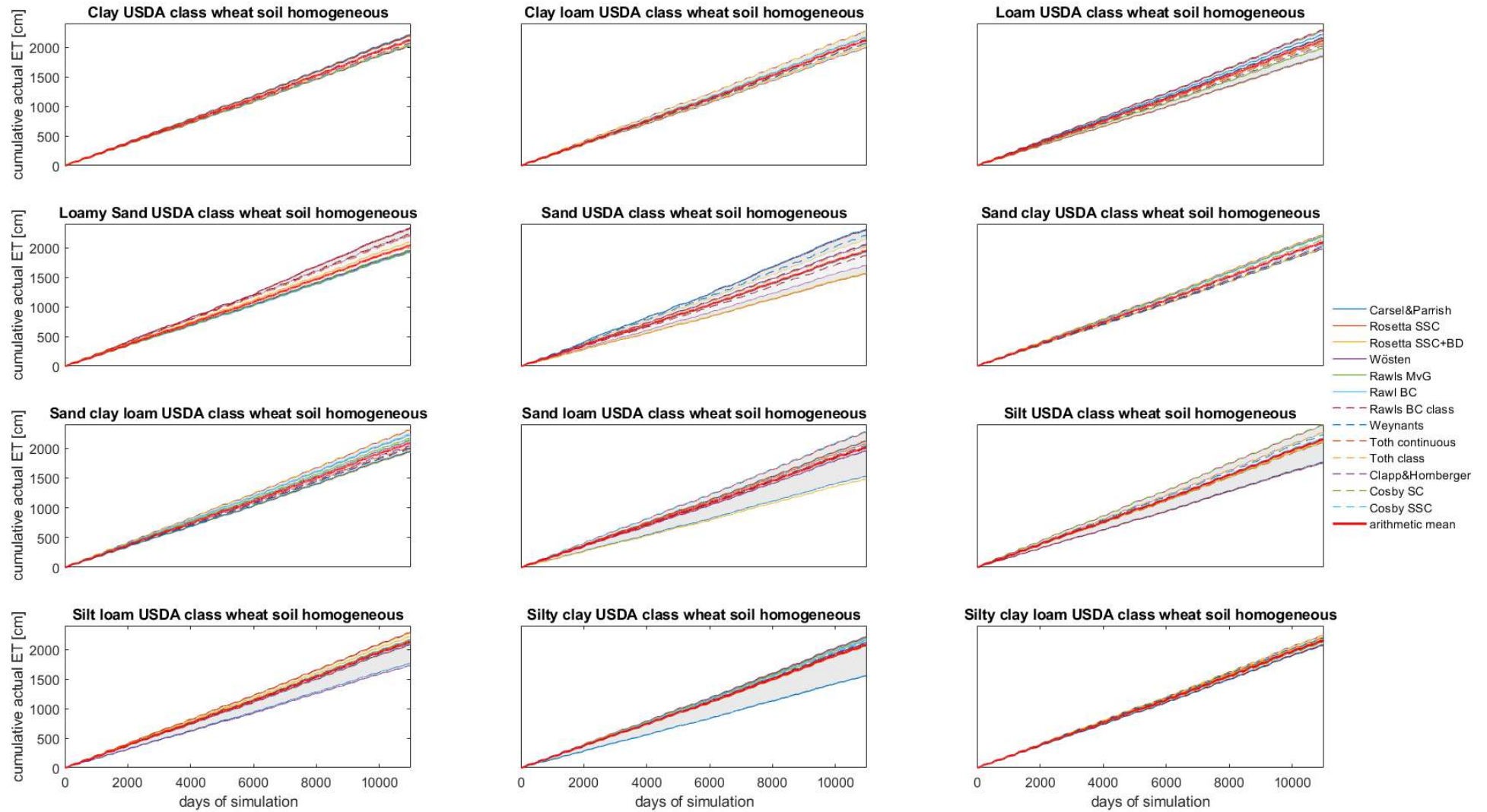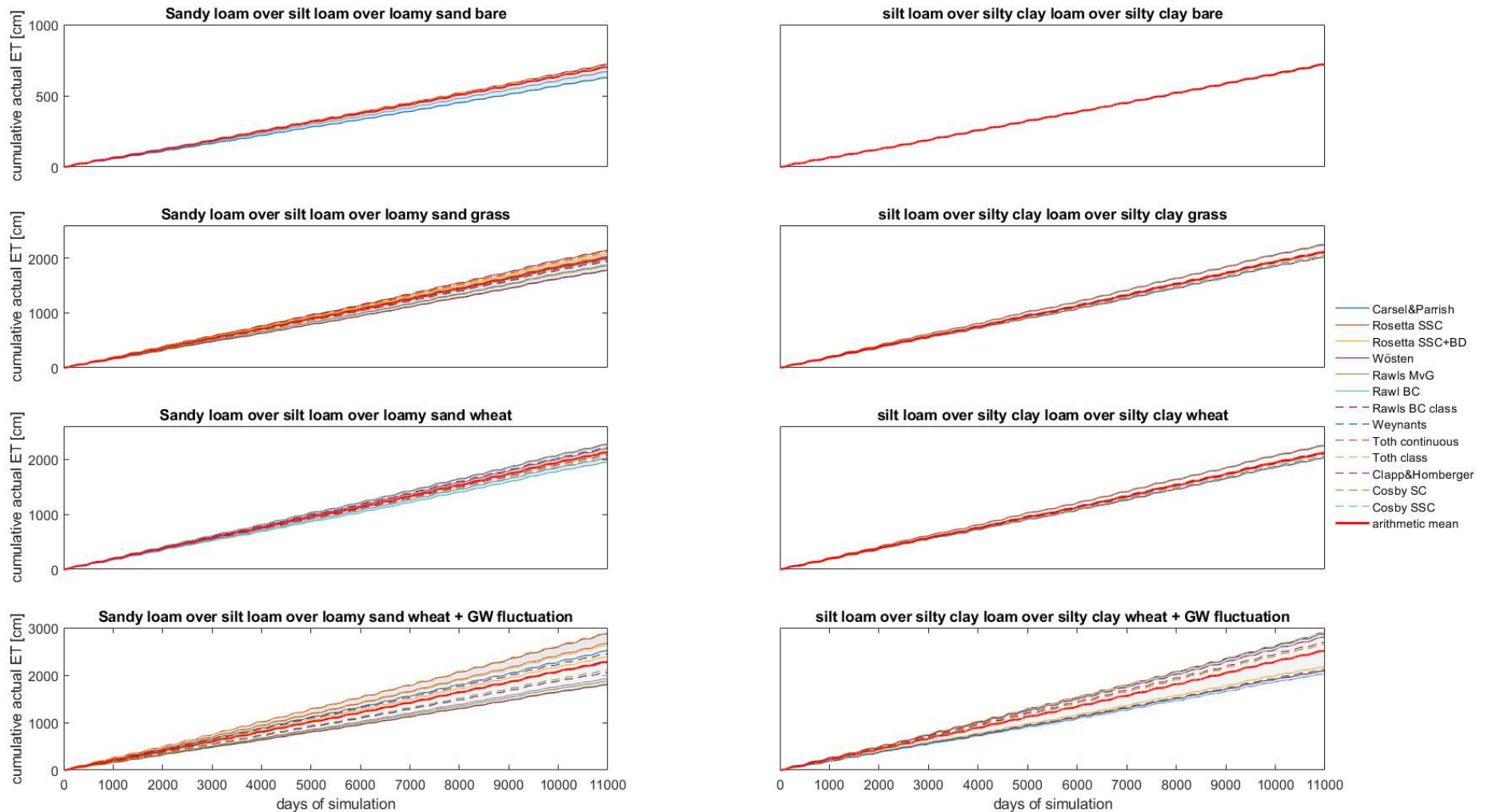
Annex Figure 2: Simulated cumulative actual evaporation E_a [cm] over the simulation period of 10988 days (30 years) for the homogeneous bare soil scenario. Light and dark grey shaded area represent the 70 and 90 % tolerance interval, respectively.
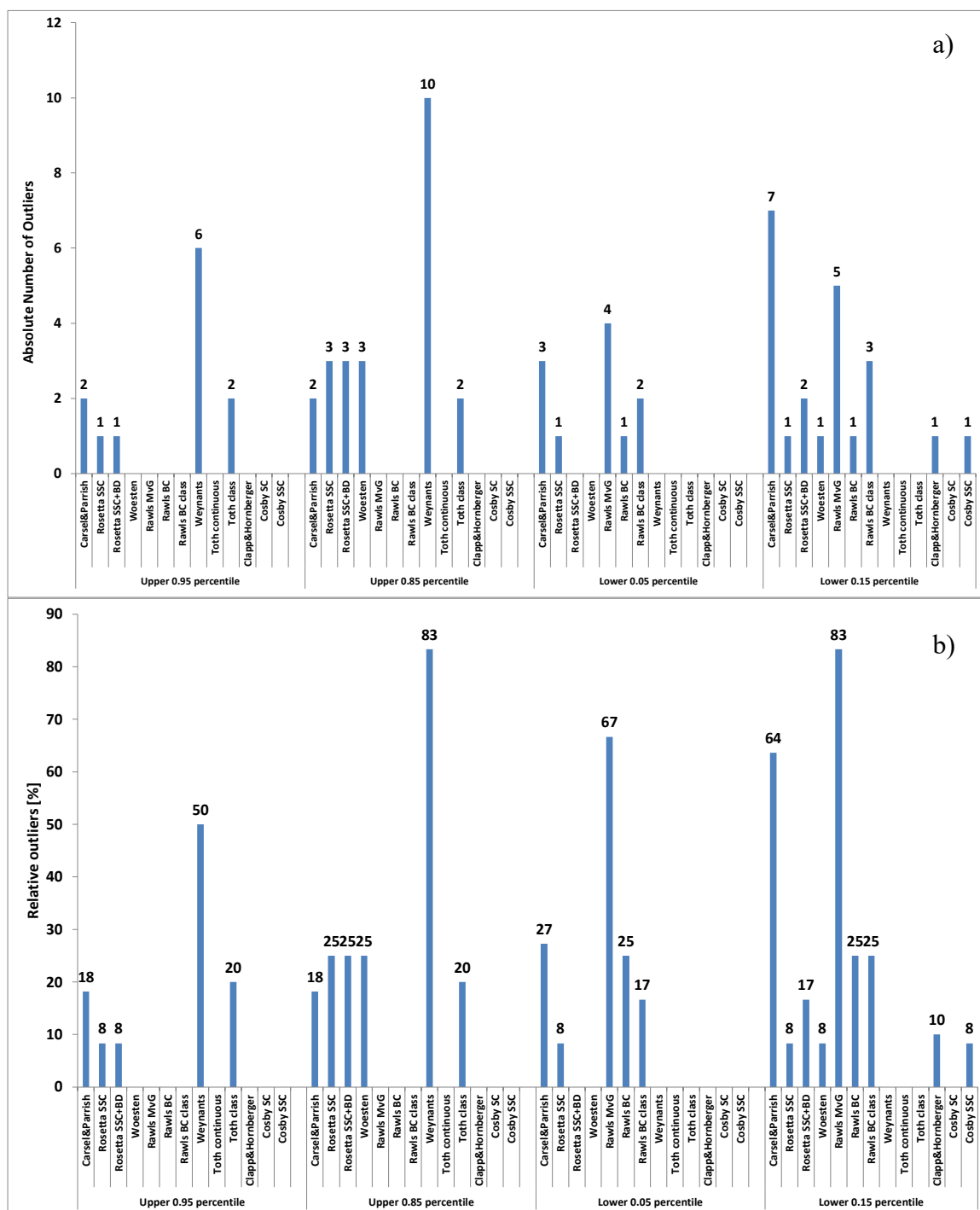
1425

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.
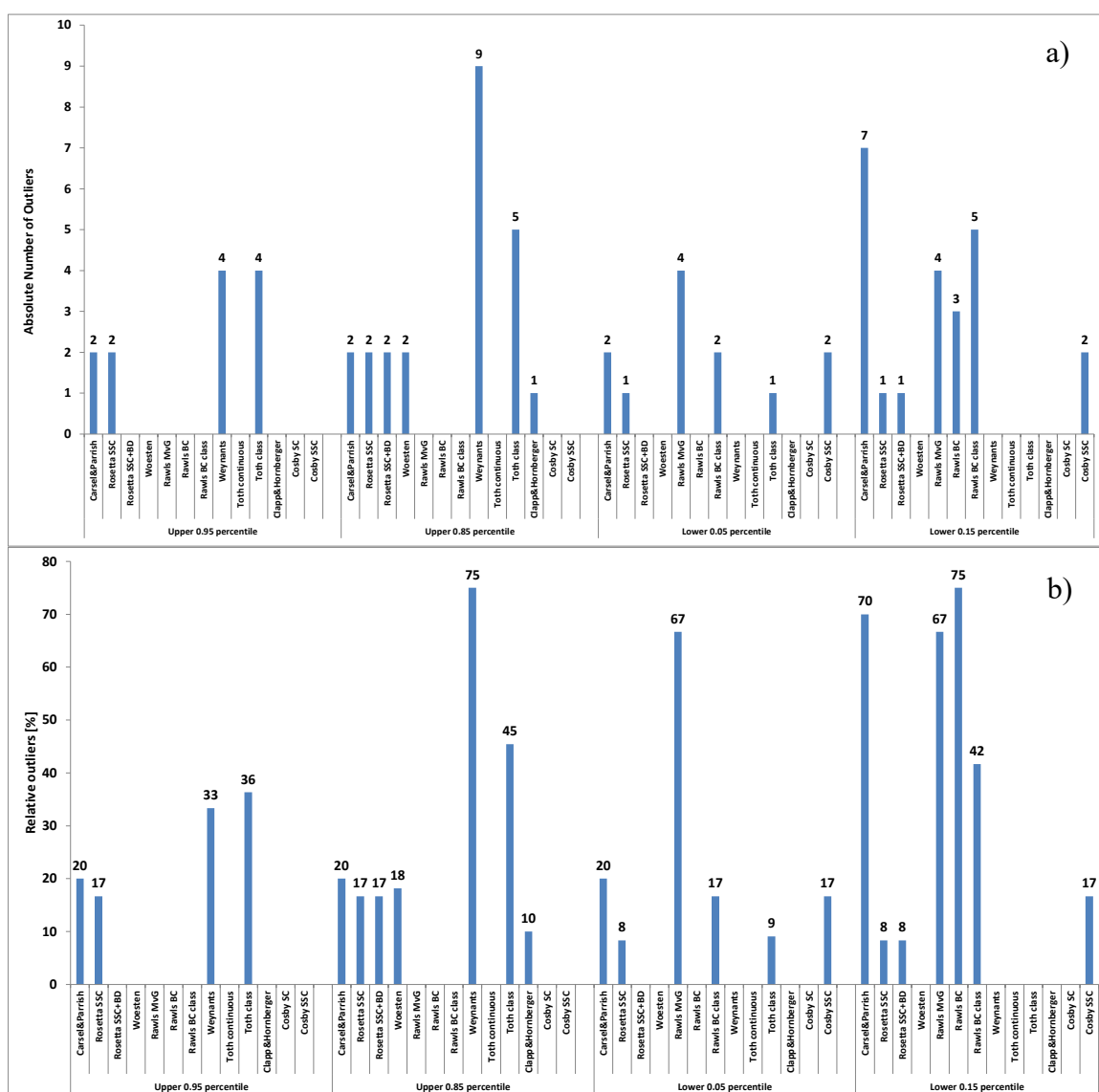
1430

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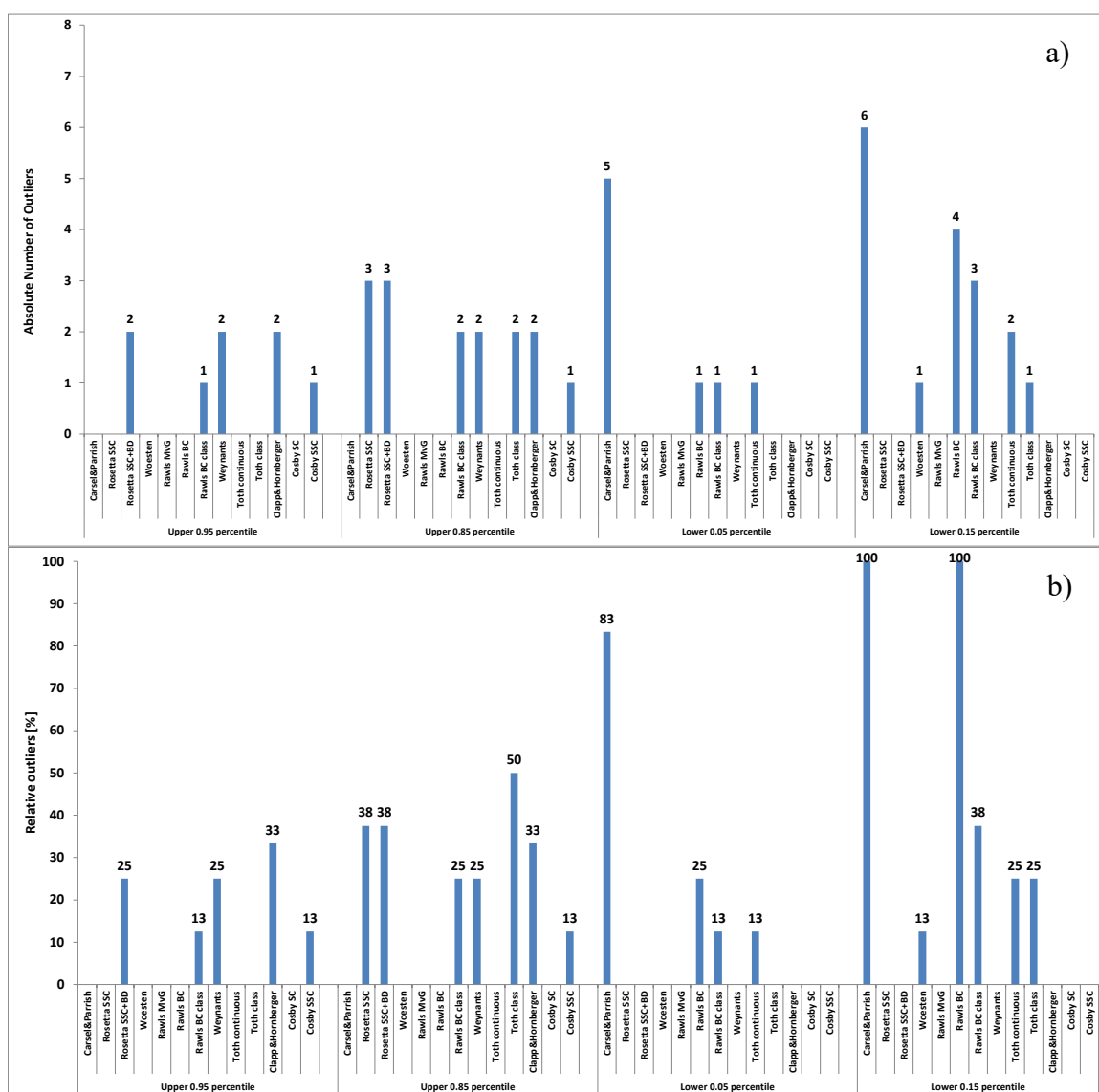
Annex Figure 6: a) Absolute and b) relative number of outliers for simulated E_a/ET_a at t_{end} for the 12 USDA soil classes (11 for Clapp&Hornberger) and the homogeneous grass vegetated soil scenario.

1445

1450

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.

1455

1460

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).