

## A framework for quantifying hydrologic effects of soil structure across scales

Sara Bonetti <sup>1,2</sup>✉, Zhongwang Wei <sup>3,4</sup> & Dani Or <sup>5,6</sup>

Earth system models use soil information to parameterize hard-to-measure soil hydraulic properties based on pedotransfer functions. However, current parameterizations rely on sample-scale information which often does not account for biologically-promoted soil structure and heterogeneities in natural landscapes, which may significantly alter infiltration-runoff and other exchange processes at larger scales. Here we propose a systematic framework to incorporate soil structure corrections into pedotransfer functions, informed by remote-sensing vegetation metrics and local soil texture, and use numerical simulations to investigate their effects on spatially distributed and areal averaged infiltration-runoff partitioning. We demonstrate that small scale soil structure features prominently alter the hydrologic response emerging at larger scales and that upscaled parameterizations must consider spatial correlations between vegetation and soil texture. The proposed framework allows the incorporation of hydrological effects of soil structure with appropriate scale considerations into contemporary pedotransfer functions used for land surface parameterization.

<sup>1</sup> Soil Physics and Land Management Group, Wageningen University, Wageningen, The Netherlands. <sup>2</sup> Institute for Sustainable Resources, University College London, London, UK. <sup>3</sup> Guangdong Province Key Laboratory for Climate Change and Natural Disaster Studies, School of Atmospheric Sciences, Sun Yat-Sen University, Guangzhou, China. <sup>4</sup> Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, China. <sup>5</sup> Department of Environmental Systems Science, Swiss Federal Institute of Technology (ETH) Zürich, Zürich, Switzerland. <sup>6</sup> Division of Hydrologic Sciences, Desert Research Institute, Reno, NV, USA. ✉email: [sara.bonetti@wur.nl](mailto:sara.bonetti@wur.nl)

Earth system and land surface models (LSMs) used for predictions of climatic and hydrological processes require spatially distributed soil hydraulic properties (SHPs) to represent surface fluxes, particularly the partitioning between infiltration and runoff<sup>1–5</sup>. SHPs are generally derived from soil pedotransfer functions (PTFs) that correlate readily-available soil information (e.g., texture, bulk density, organic matter) with difficult-to-measure soil hydraulic parameters<sup>6–10</sup>. In practice, PTFs are trained using measurements performed on laboratory-scale samples obtained from relatively uniform soils (i.e., agricultural fields)<sup>11,12</sup>. This sampling and measurement bias underrepresents the dominant role that soil structure (i.e., aggregation of soil particles, biopores formed by plant roots and earthworms) exerts on surface fluxes<sup>1,2,9,12–15</sup>. Effects of soil structure are expected to dominate the hydrological response in natural (non-arable) lands such as forested landscapes ( $\approx 30\%$  of the Earth's land surface<sup>16</sup>) and in regions with intense rainfall rates.

Furthermore, the standard parameterization of SHPs in LSMs uses sample scale information, often with little consideration of effective (scale informed) parameter values that account for the heterogeneity (in space and time) of natural landscapes and emergent hydrological nonlinearities and feedbacks<sup>12,17–24</sup>. To this end, a definition of functional relationships between landscape attributes and SHPs, in combination with appropriate regionalization functions and scaling operators<sup>22,24</sup>, might provide a way forward for a systematic and physically-based definition of SHPs across scales and locations<sup>21,23–25</sup>.

Focusing primarily on biologically-enhanced soil structure, studies have shown consistent links between vegetation as a promoter of biological activity and soil structure, which, in turn, affects soil hydrological response and properties<sup>26–31</sup>. It has been shown that the hydrological response of vegetated landscapes may vary significantly from non-vegetated areas in terms of runoff-infiltration partitioning under similar climatic conditions<sup>14,31–34</sup>. In their seminal work, Dunne and co-authors<sup>33</sup> have shown strong positive feedbacks between the fraction of vegetation cover and average infiltration, as illustrated in Fig. 1i where the effective infiltration for different rainfall intensities (after accounting for canopy interception) varies with vegetation cover and affects runoff generation. A meta-analysis<sup>26</sup> has shown that the infiltration capacity in arid landscapes increases as a power-law function of above ground vegetation biomass, whereas the correlation was less significant in humid climates. Across a climatic gradient from xeric to hydric, the slope of the power-law relationship between above ground biomass and infiltration capacity decreased, with no discerned effects of soil texture on the relationship. Similarly, positive relationships between plant Leaf Area Index (LAI) or vegetation biomass<sup>27,30</sup> and saturated soil hydraulic conductivity ( $K_s$ ) have been observed, with stronger vegetation effects manifested in fine-textured soils compared to coarse soils<sup>9,30</sup>. This nuanced and important aspect of soil texture-dependent structure effects on hydrological response may be misinterpreted under certain conditions as elaborated shortly. The inherent links between vegetation, biological activity, and soil structure development and maintenance imply an increase in transport properties (prominently the soil hydraulic conductivity) in structured soils relative to repacked samples or those obtained from tilled soil. We hypothesize that persistent bias in SHPs of vegetated lands could be rectified by considering effects of vegetation on spatially-informed PTFs.

The study was motivated by the ever-increasing need for spatially-resolved soil parameters and by the limitations of currently used PTFs which omit soil structure and lack scale information<sup>2,5,15,30,35</sup>. We capitalize on highly resolved soil and vegetation cover maps to simultaneously address these limitations

of PTFs and propose a framework for inclusion of soil structure effects informed by vegetation and soil patterns at different scales. For simplicity, in this study we ignore the relatively limited effects of abiotic processes (shrinking cracks and freeze-thaw) to focus on the primary driver of soil structure, namely vegetation-promoted biological activity. We propose a simple framework for parameterization of soil structure effects on SHPs via observable vegetation indices (as illustrated in Fig. 1) with emphasis on how macroporosities alter the saturated hydraulic conductivity and infiltration response of landscapes. To maintain our focus on improving land surface parameterization, we first use a minimalist infiltration model and synthetic examples to disentangle the nonlinear feedbacks between vegetation cover, soil texture, and climate and evaluate the conditions and scales where soil structure corrections are likely to have a major impact. In addition, while a systematic evaluation of the proposed parameterization is hindered by the limited observability of soil structure effects at large spatial scales<sup>15</sup>, catchment-scale simulations with the Noah-MP LSM are used to highlight conditions in which soil structure corrections may strongly alter the local and catchment-scale hydrologic response. Specifically, we will show that, under certain soil and climatic conditions (i.e., fine-textured soils with dense vegetation cover subjected to high-intensity rainfall events), small scale soil structure plays a key role on the dynamics and timing of the hydrologic response even at the large scales relevant to LSMs.

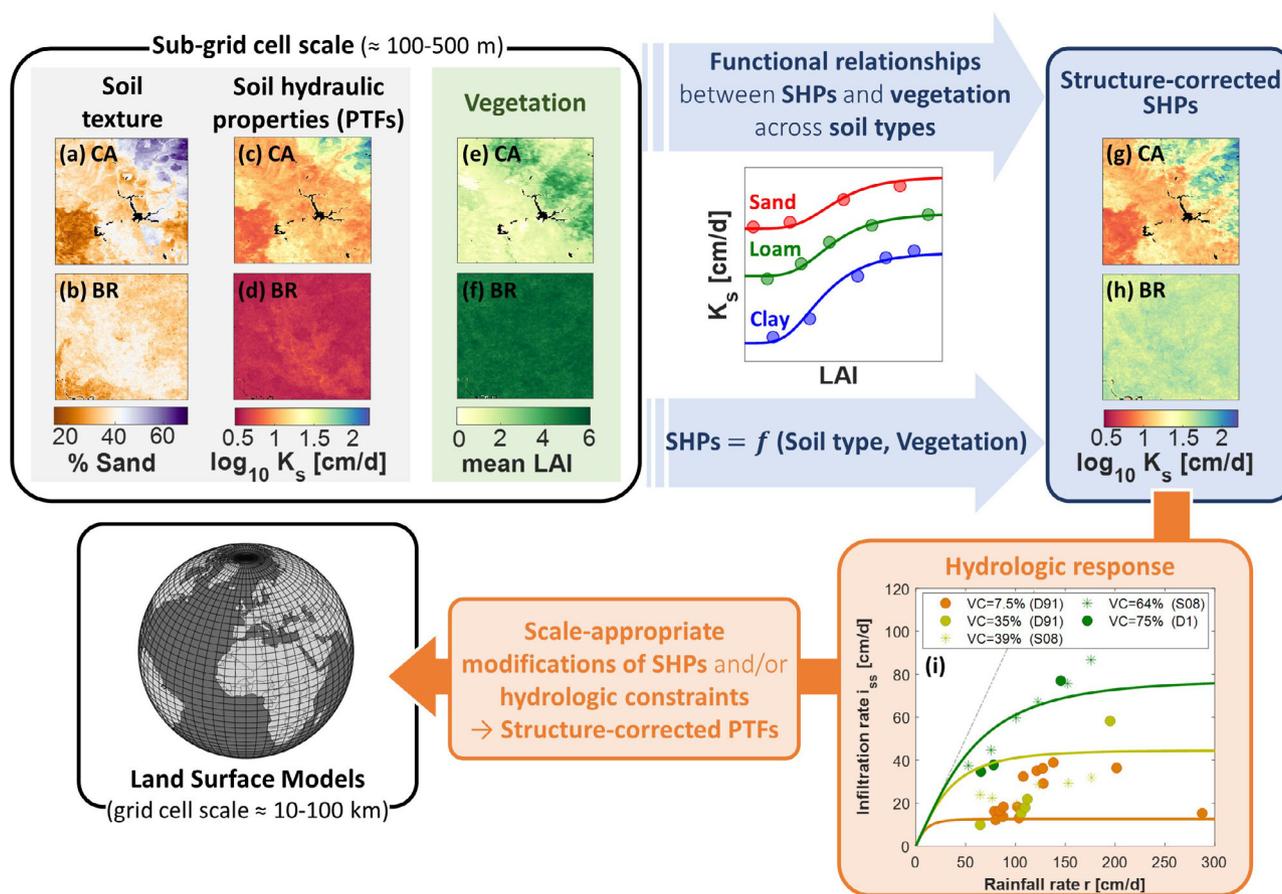
## Results and discussion

**Soil structure characterization using vegetation indices.** To harness the availability of spatially resolved soil maps and vegetation metrics, we formalize the links between vegetation and soil structure and their potential effects on surface hydrology. In the analyses we consider the primary effect of biologically-induced soil macroporosity (e.g., presence of roots) on modifying the soil hydraulic conductivity near-saturated conditions<sup>36</sup> (see Fig. 2a), while the effects on other parameters of the soil retention curve are deemed to be relatively minor<sup>13</sup>. This allows us to focus on soil structure modifications of soil saturated hydraulic conductivity (relative to its value based on texture only) and associated landscape hydrological responses such as rainfall partitioning to infiltration and runoff.

We evaluated vegetation indices including aboveground vegetation biomass,  $B$  [ $\text{kg m}^{-2}$ ], and LAI [ $\text{m}^2_{\text{leaf}} \text{m}^{-2}_{\text{ground}}$ ]. Experimental data<sup>26,27,30</sup> (see Fig. 2b and Supplementary Fig. S1) show a systematic increase in structure-modified saturated hydraulic conductivity (denoted as  $K_s$ ) with increasing values of vegetation attributes (biomass or LAI). For practical implementation without full mechanistic modeling of vegetation effects on soil structure, we propose the following empirical relation for approximating structure effects on the soil saturated hydraulic conductivity,

$$K_s = K_{s,\max} - \frac{K_{s,\max} - K_0}{1 + \left(\frac{V}{\alpha}\right)^\beta}, \quad (1)$$

where  $V$  is a vegetation metric (i.e.,  $B$  or LAI),  $\alpha$  and  $\beta$  are shape parameters,  $K_0$  is the saturated hydraulic conductivity of texture-only or unstructured soil (i.e., when  $V = 0$ ), and  $K_{s,\max}$  is the maximum saturated hydraulic conductivity of a soil with fully developed structure (i.e., fully vegetated surface). The postulated sigmoidal relation in Eq. (1) was fitted to literature data in Fig. 2b and Supplementary Fig. S1, and values of the shape parameters  $\alpha$  and  $\beta$  were defined based on the data comparison (here  $\alpha = 0.4$ ,  $\beta = 4$  for  $V = B$ , and  $\alpha = 4.5$ ,  $\beta = 5$  for  $V = \text{LAI}$ ). The extreme values of  $K_s$  for soil with no structure and the same soil with well-developed structure (marked by  $K_0$  and  $K_{s,\max}$  respectively) vary with soil textural class. While  $K_0$  is easily obtained from traditional texture-based PTFs, the magnitude of  $K_{s,\max}$  is derived here based



**Fig. 1 Proposed methodology for the development of structure-corrected soil hydraulic properties to be used in Land Surface Models.** Functional relationships between soil hydraulic properties (SHPs) and vegetation cover across different soil types and biomes will be first defined, thus allowing the characterization of soil structure effects on SHPs using vegetation metrics as surrogates. Such functional relationship can be directly employed to correct SHPs from traditional pedotransfer functions (PTFs). The effect of soil structure on the hydrological response (in terms of runoff-infiltration partitioning) is then quantified from point to grid cell scale to estimate scale appropriate modifications of SHPs accounting for the effect of spatial heterogeneities on the areal-averaged hydrological response. Spatial distribution of (e, f) mean LAI (MODIS data, 500 m resolution, year 2017), a, b percentage of sand at 30 cm depth (SoilGrids Data, 250 m resolution), c, d saturated hydraulic conductivity ( $K_0$ ) obtained from soil textural properties using PTFs from Saxton<sup>59</sup>, g, h structure-corrected saturated hydraulic conductivity ( $K_s$ ) computed by means of Eqs. (1) and (2) with  $V = LAI$  (mean annual value),  $\alpha = 4.5$ ,  $\beta = 5$ . CA and BR refer to synthetic case studies in California and Brazil, respectively; black represents no-data points. i The effective infiltration rate as a function of rainfall for different levels of vegetation cover: symbols represent data digitized from Dunne et al.<sup>33</sup> (circles) and Stone et al.<sup>34</sup> (asterisks), solid lines are obtained from the theoretical areal-averaged infiltration rate (see Supplementary Methods) assuming  $m_{K_0} = 34 \text{ cm day}^{-1}$ ,  $v_{K_0} = 900 \text{ cm}^2 \text{ day}^{-2}$ , and 45% sand. Line colors refer to different amounts of aboveground vegetation biomass  $B$  (orange =  $0.15 \text{ kg m}^{-2}$ , yellow =  $0.4 \text{ kg m}^{-2}$ , green =  $1.4 \text{ kg m}^{-2}$ ). The gridded globe was retrieved from <https://commons.wikimedia.org> under the CC BY-SA 3.0 license and modified to grayscale.

on soil type, as detailed below. The postulated relationship between vegetation indices and structure-modified saturated hydraulic conductivity represents a vast simplification of other influences such as vegetation type and details of soil biological activity, as well as influences of abiotic processes (that were neglected here for simplicity).

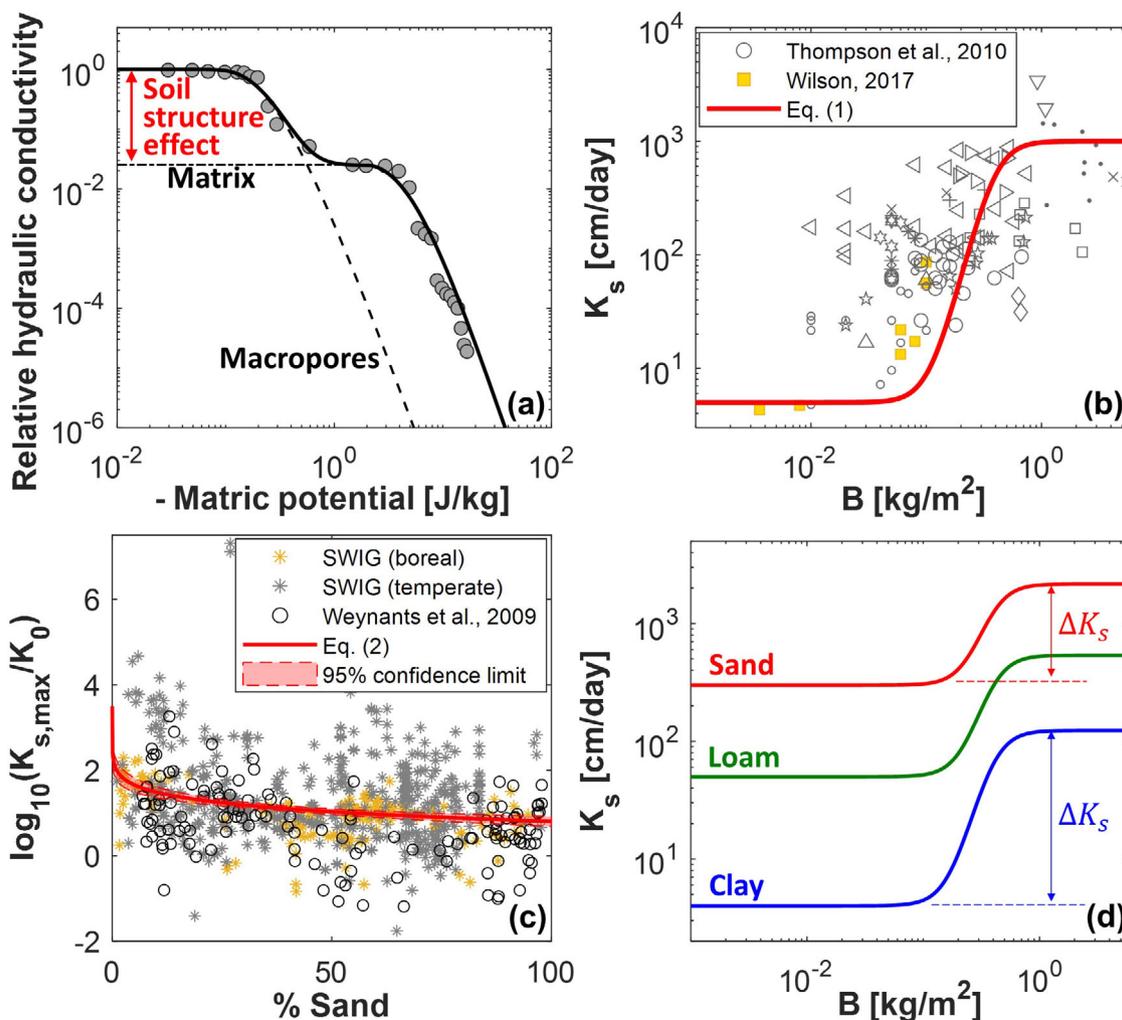
As noted earlier, the degree by which vegetation influences the soil saturated hydraulic conductivity (i.e., the value of  $K_{s,max}$ ) varies with soil texture<sup>9,30</sup>. Results from previous work<sup>9</sup> show that the effect of soil structure on saturated conductivity can be expressed by the ratio of saturated conductivity of soil with structures ( $K_{s,max}$ ) to saturated matrix flow ( $K_0$ ) and that soil structural effects are larger in finer textured media soils. The ratio of structured to matrix saturated conductivities was further evaluated for a series of observations in temperate and boreal regions retrieved from the SWIG database<sup>37</sup>, confirming the more pronounced role of soil structure for finer texture soils (Fig. 2c). Based on these observations, we postulate that the ratio of  $K_{s,max}$  to  $K_0$  will decrease in proportion to the soil's percentage of sand

fraction  $Sa$  [%] according to the relation

$$\log_{10} \frac{K_{s,max}}{K_0} = 3.5 - 1.5 \cdot Sa^{0.13} \tag{2}$$

where the coefficients are derived from a fitting procedure (see Fig. 2c). Specifically, the regression was performed using the Matlab nonlinear fitting tool and constrained so that the upper limit at 0% sand content is equal to the 90 percentile of the  $\log_{10} \frac{K_{s,max}}{K_0}$  values of measurements in soil samples having sand content between 0 and 10%.

Eqs. (1) and (2) enable observation-driven characterization of the impact of increasing vegetation cover on the soil saturated hydraulic conductivity across different soil textures, as exemplified in Fig. 2d where  $K_s$  is depicted as a function of above-ground vegetation biomass for typical sandy, loamy, and clayey soils: while soil structure only partially modifies  $K_s$  for sandy soils, the saturated hydraulic conductivity can change by 1 to 2 orders of magnitudes for finer texture soils (e.g., clay). This highlights how the proper quantification of soil structure effects is not merely



**Fig. 2** Soil-structure characterization via vegetation indices and across soil textures. **a** Model of structured porous media applied for aggregated loam where soil matrix and macroporosities differently affect the hydraulic conductivity functions (modified from Tuller and Or (2002)<sup>36</sup>). **b** Saturated hydraulic conductivity as a function of above-ground vegetation biomass (gray symbols are xeric sites from Thompson et al.<sup>26</sup>, yellow squares are results from a rainfall simulator in Sardinia, Italy<sup>27</sup>). The red line is given by Eq. (1) with  $V = B$ ,  $\alpha = 0.4$ ,  $\beta = 4$ ,  $K_0 = 5 \text{ cm d}^{-1}$ , and  $K_{s,\text{max}} = 1000 \text{ cm d}^{-1}$ . **c** Relationship between percentage sand and  $\log_{10} K_{s,\text{max}}/K_0$  (black circles are data from Weynants et al. (2009)<sup>9</sup>, asterisks are data for temperate and boreal regions of the SWIG database<sup>37</sup>, the red line is Eq. (2)). **d**  $K_s$  as a function of aboveground vegetation biomass for typical sand (85% sand,  $K_0 = 300 \text{ cm d}^{-1}$ ), loam (50% sand,  $K_0 = 50 \text{ cm d}^{-1}$ ), and clay (5% sand,  $K_0 = 4 \text{ cm d}^{-1}$ ) soils, computed by means of Eqs. (1) and (2) with  $V = B$ ,  $\alpha = 0.4$ , and  $\beta = 4$ .

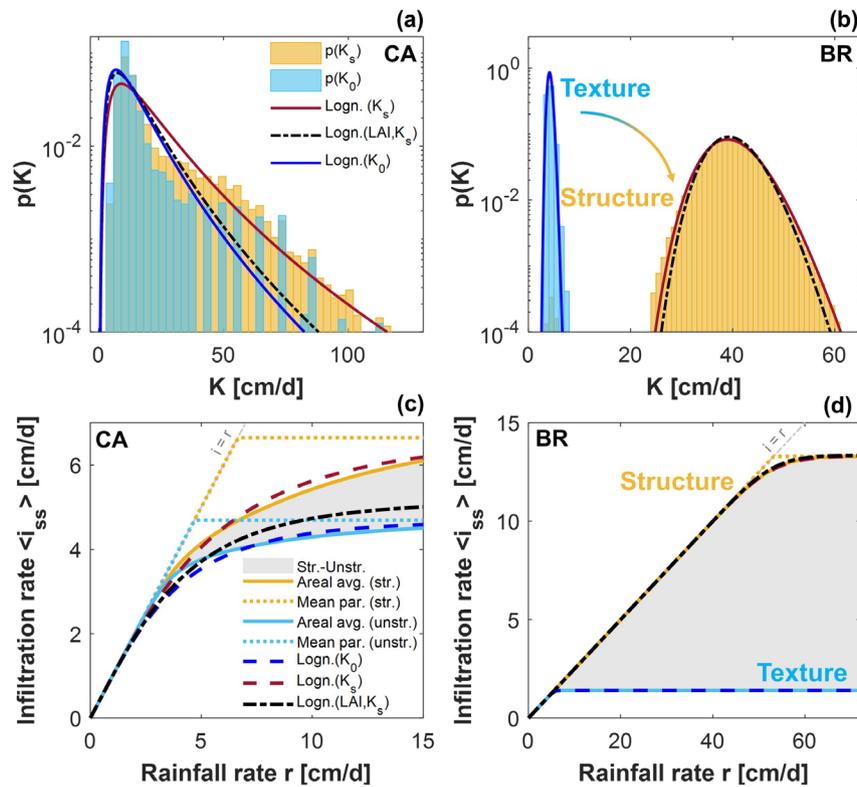
related to the amount of vegetation/biological activity, but the magnitude of its effect is non-linearly tied to the type of soil under consideration.

**Effects of soil structure on the hydrological response.** To quantify how soil structure affects the hydrological response of vegetated landscapes, we first used a simplified infiltration model based on the two-term Philip's equation<sup>38</sup> with Brooks and Corey parameterization of the SHPs (see Supplementary Methods for details). Vegetation-enhanced soil structure effects were accounted for by modifying the saturated hydraulic conductivity according to Eqs. (1) and (2). We focus on the landscape hydrological response at later stages of a rainfall event, when the impact of capillary terms has diminished (i.e., sorptivity affects initial infiltration in a similar manner for unstructured and structured soils), and infiltration rate is controlled by gravity and saturated hydraulic conductivity. This is also the stage where surface runoff becomes important, hence of interest for this study. Representative results for a point-infiltration process into typical profiles of sand, loam, and clay soils are shown in Supplementary

Fig. S2. Generally, soil structure-correction delays the time of ponding and reduces surface runoff due to an increase in the soil infiltration capacity. The effect of soil structure on infiltration metrics is more pronounced in fine-textured soils under most rainfall rates, whereas the impact on hydrological flux partitioning in coarse soils becomes relevant only at very high rainfall rates (Supplementary Fig. S2).

#### Effects of spatial organization of vegetation and soil texture.

The results in the previous section were mostly diagnostic, focusing on the behavior at the point scale. To include realistic landscapes at scales relevant to land surface modeling (i.e., a catchment or a LSM grid cell) we now investigate how heterogeneities and spatial variations in soil texture and vegetation cover affect the (subgrid) magnitudes and spatial arrangements of the modified soil hydraulic properties. Specifically, to quantify an areal-averaged response of larger domains considering vegetation and soil heterogeneities and evaluate the impact of subgrid spatial arrangements, we derived an expression for the areal-averaged infiltration rate,  $\langle i_{ss} \rangle$ , when capillary influence diminishes and



**Fig. 3** Soil structure effect on areal average infiltration rate. **a, b** Probability distributions of unstructured ( $K_0$ , blue bars) and structured ( $K_s$ , yellow bars) saturated hydraulic conductivities values shown in the maps of Fig. 1: solid lines show the fitted lognormal distributions (red = structured, blue = unstructured). **c, d** Steady-state infiltration rate as a function of rainfall rate: solid lines show areal averaged quantities while dotted lines are results obtained assuming homogeneous domains with mean parameter values (yellow = structured, blue = unstructured). The dashed lines are theoretical areal-averages obtained from Eq. (3). The black dash-dotted lines show pdfs and theoretical results obtained by modifying the unstructured distribution using mean LAI and percentage sand values, see Eq. (S8) in the Supplementary Methods. Left and right panels refer to results for the regions in CA and BR, respectively.

infiltration rate attains near steady-state conditions (see Supplementary Methods for details). The analytical form of the steady-state areal average assumes a lognormal distribution of soil saturated hydraulic conductivity as supported by experimental evidence<sup>39–42</sup> (see also comparison with data in Fig. 3a, b) and reads

$$\langle i_{ss} \rangle = \frac{r}{2} + \frac{r}{2} \operatorname{erf} \left( \frac{\mu + \ln \frac{1}{3r}}{\sqrt{2}\sigma} \right) + \frac{e^{\mu + \frac{\sigma^2}{2}} \operatorname{erfc} \left( \frac{\mu + \sigma^2 - \ln 3r}{\sqrt{2}\sigma} \right)}{6} \quad (3)$$

where  $r$  is the rainfall rate (assumed homogeneous in space and during the rainfall event in the applications here), and the parameters  $\mu$  and  $\sigma$  are related to the mean and variance of the probability density function (pdf) of  $K_s$ . Among the many simplifying assumptions, we neglect here effects of topography and lateral water flows either above or below ground, consider short time windows (i.e., a rainfall event), and do not distinguish between different runoff generation mechanisms. These simplifications are motivated by our objective of using a minimalist framework as a benchmark for testing the salient features of the proposed soil-structure modified SHP parameterization and impacts on runoff-infiltration partitioning across combinations of soil textures, rainfall amounts, and vegetation cover. The behavior of the steady-state infiltration rate for a loamy soil and different amounts of aboveground vegetation biomass is shown in Fig. 1i, displaying a decrease in runoff as vegetation increases, in line with previous observations<sup>33,34</sup> (also shown in Fig. 1i) and thus supporting the use of vegetation indices as surrogates for soil-structure parameterization. Furthermore, accounting for the

spatial distribution of saturated hydraulic conductivities (instead of using mean parameter values, as for the point-infiltration process in Supplementary Fig. S2) was instrumental in recovering the observed gradual increase of infiltration rates with increasing rainfall intensity. Such gradual increase is associated with the spatial distribution of  $K_s$  over the considered domain: the higher the variance of the  $K_s$  distribution, the more gradual is the increase in infiltration as a function of rainfall (Supplementary Fig. S3). Conversely, changes in the mean  $K_s$  value shift the response curve up and down (i.e., change the magnitude of the infiltration rate) without modifying its overall shape as a function of rainfall rate (Supplementary Fig. S3).

We further examine the integrated runoff-infiltration response at a catchment scale using a typical LSM grid cell (generally between  $10^2$  and  $10^4$  km<sup>2</sup>). At these large scales, remotely sensed vegetation indices are particularly useful for quantifying the spatial distribution and density of vegetation and the associated impact on SHPs. For simplicity, in this case we opted for using mean annual LAI as a surrogate for soil structure modifications, noting that other metrics could be equivalently suitable (e.g., gross primary productivity) provided that a relationship between the altered SHP (i.e.,  $K_s$ ) and the vegetation variable of interest has been established (i.e., Eq. (1)). The use of a constant value of LAI (i.e., neglecting seasonal variations) is motivated by the fact that seasonal changes in LAI do not affect the dominant root system or below-ground biomass distribution (which are likely to be more stable throughout the year for a natural vegetation cover).

For the evaluation of the proposed SHP parameterization at the LSM grid cell we considered two regions of approximately 80 km

by 80 km in Northern California (CA) and in the Brazilian Amazon (BR), and used MODIS<sup>43,44</sup> and SoilGrids<sup>45</sup> data for the spatial distribution of *LAI* and soil properties, respectively (see Fig. 1 and Supplementary Fig. S4). These two case studies were chosen as diagnostic experiments to illustrate the effect of the proposed parameterization over landscapes characterized by different soil and vegetation cover and provide general guidance for what scale-appropriate modifications of SHPs might mean. We emphasize that the focus of the analyses is not the hydrological model performance, but we seek to highlight the influence of the two parameterization schemes via their effect on infiltration-runoff response for different spatial arrangements of vegetation and soil texture. This will further allow us to identify strategies (and constraints) for the definition of scale-appropriate soil hydraulic parameterizations (as discussed further below). The CA grid cell is characterized by a soil texture gradient from clay to sand with patchy vegetation cover concentrated primarily in loamy areas, whereas the grid cell in Brazil (BR) is mostly clayey with a uniform and dense vegetation cover (see Fig. 1). The spatial distribution of texture-based saturated hydraulic conductivity ( $K_0$ ) and its structure-corrected counterpart ( $K_s$ ) are shown in Fig. 1. While the saturated hydraulic conductivity was slightly modified in the vegetated patches of the CA domain, its value is changed up to two orders of magnitude over the BR area due to the combination of high vegetation cover and fine texture soil.

The areal average behavior of the system for texture-based and structure-corrected SHP conditions was quantified by solving Philip's equation at every point within the grid cell (i.e., over sub-grid areas of 250 m by 250 m, a resolution imposed by the SoilGrids database) for increasing rainfall rates. Results are shown in Fig. 3 and Supplementary Figs. S5 and S6. As expected, the BR domain experiences faster and more extensive ponding compared to the CA case for conditions where soil structure is neglected (due to the clayey nature of the soil in BR). However, accounting for soil structure modifies the average behavior of the BR grid cell rendering it more similar to the CA domain, with delayed and lower surface ponding for the same rainfall rates (between approximately 1 and 20 cm d<sup>-1</sup>) and initial conditions (Supplementary Figs. S5 and S6). These results highlight the important role of soil structure in the magnitude and timing of hydrological response. The areal averaged infiltration rate as a function of rainfall at steady state is shown in Fig. 3c, d (solid lines), stressing the bias in runoff estimate if soil structure is neglected. In particular, while such bias is negligible for low rainfall rates where no runoff occurs, it becomes significant as rainfall intensity and duration increase. Further comparison with the solution obtained considering a homogeneous grid cell with mean parameter values (dotted lines in Fig. 3c, d) shows that accounting for spatial heterogeneities allows us to recover the gradual increase of infiltration with rainfall<sup>33,34,46</sup>. The analytical solution for the areal-averaged infiltration rate (Eq. (3), dashed lines in Fig. 3c, d) is in good agreement with the numerical areal average.

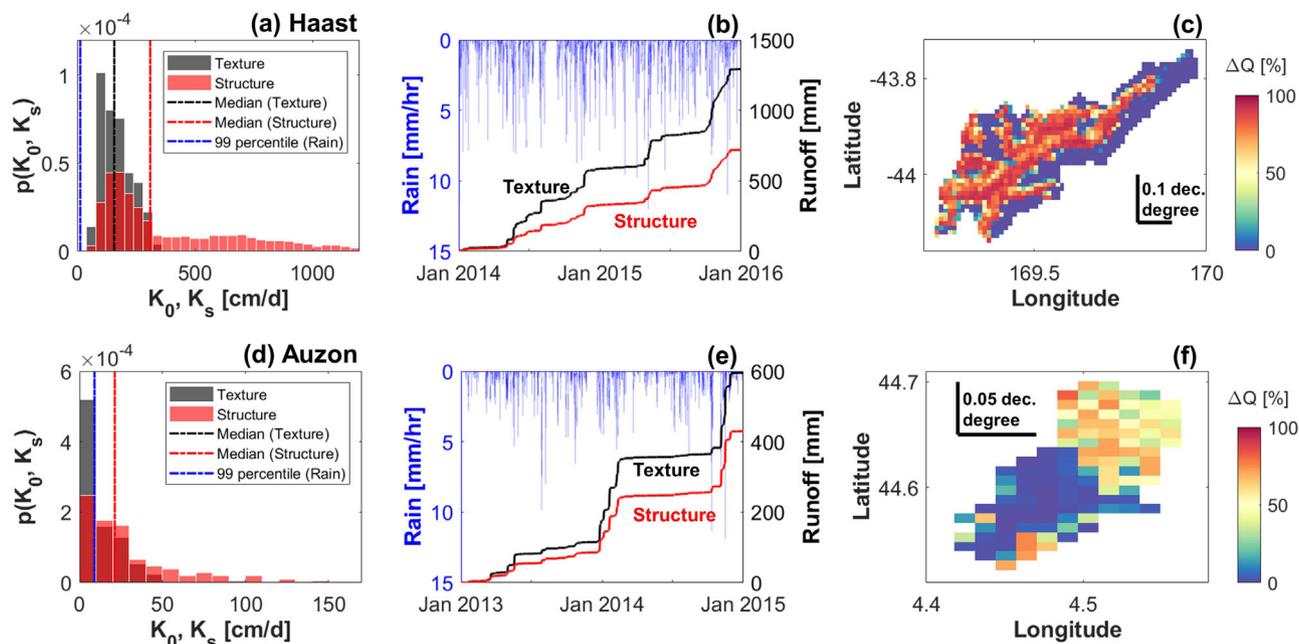
The expression for the areal-averaged steady-state infiltration rate given by Eq. (3) requires the knowledge of the pdf of the spatially distributed structure-corrected saturated hydraulic conductivity. This may be obtained via a point-by-point modification of the original soil saturated hydraulic conductivity (based on Eqs. (1) and (2)) at the sub-grid cell scale to define the parameters of the pdf of  $K_s$ . Fortunately, this strict requirement could be relaxed in the absence of cross-correlation between soil texture and vegetation cover (as suggested from a preliminary analysis for natural vegetation across different biomes—see Supplementary Fig. S10). Consequently, the pdf of  $K_s$  is derived from the pdf of  $K_0$  modified by mean values of soil texture (% sand) and vegetation cover (*LAI*) to allow for analytical representation (see

Supplementary Methods for details). The resulting steady-state infiltration rate is shown in Fig. 3c, d (black dash-dotted line), exhibiting good agreement with the numerical (point-by-point evaluation) areal-average for the BR case, while it underestimates the effect of soil structure for the CA domain. Within this simplified analysis, the lack of vegetation cover correlation with soil texture for the Brazilian case permits analytical representation of the areal response using mean surrogate properties, which would not be possible in case of cross-correlation (as in the CA case where vegetation is more abundant over intermediate and coarser soils, see Supplementary Fig. S7). In areas where significant correlation exists between soil texture and vegetation, a numerical grid-specific parameterization would be needed (see also the synthetic simulations presented in Supplementary Figs. S8 and S9 and discussed in the Supplementary Results). The sensitivity of the outcome to cross-correlation is important for other upscaling methods that attempt to derive effective spatial properties influenced by multiple variables (such as the potential of vegetation-soil cross-correlation).

**LSM application: evaluating conditions for soil structure activation.** Notwithstanding the physical basis and simplicity of the approach for considering vegetation effects on soil structure and alteration of SHPs, in the following we address the important question: do these alterations actually matter at such large scale applications? In fact, for a range of conditions (i.e., low rainfall rates, sandy top soils) the soil structural macroporosities may be unimportant or not activated under most conditions<sup>15</sup>. We thus opted for two simple motivating examples and performed 2-year LSM simulations to evaluate the hydrological response of two contrasting catchments, namely the Haast river catchment located on the West Coast of the South Island of New Zealand (NZ) and a Mediterranean catchment (Auzon) of the Ardèche region, France.

The Haast catchment (1026 km<sup>2</sup>) is characterized by a mean annual rainfall of up to approximately 6000 mm yr<sup>-1</sup> and elevations between 90 and 2245 m a.s.l.<sup>47–49</sup>, with rather coarse-textured soils (Supplementary Fig. S11) and a vegetation cover ranging from podocarp-dominated forests in the lowlands, hardwood and broadleaf forests in the montane region, small tree conifers present in the subalpine zone, and mostly shrublands and grasslands at higher elevations<sup>50</sup>. The Auzon catchment<sup>51</sup> (116 km<sup>2</sup>) has elevations ranging between 140 and 1019 m a.s.l., average yearly rainfall of approximately 850–900 mm throughout the basin (which is an intermediate value between the 500 mm annual average of the Rhône Valley to the east and the 2000 mm of the Ardèche Mountains to the west), and mainly clay and fine silt soils (Supplementary Fig. S11). Approximately 27% of the Auzon catchment consists of forest, while pastures, vineyards, and sparsely vegetated areas represent the 17, 19, and 14 % of the basin, respectively (the remaining land is occupied by crops, natural grassland, and urban areas). Figure 4a and d shows the pdf of the texture-based ( $K_0$ ) and structure-corrected ( $K_s$ ) soil saturated hydraulic conductivity for the two catchments. The rather coarse texture of the Haast basin provides relatively high  $K_0$  values, with a median of 157.4 cm d<sup>-1</sup> that is increased to 310.1 cm d<sup>-1</sup> once vegetation-based structure corrections are introduced (Fig. 4a). Conversely, the finer texture soils of the Auzon catchment result in lower  $K_0$  values (median  $K_0 = 9$  cm d<sup>-1</sup>) that are increased once structure-corrections are considered (median  $K_s = 21.3$  cm d<sup>-1</sup>, Fig. 4d).

Results from 2-year simulations with the Noah-MP LSM show a marked impact of soil structure on the spatial distribution of mean annual runoff values as well as cumulative runoff for both catchments (Fig. 4), with surface runoff being reduced by up to



**Fig. 4** Effect of soil structure informed soil hydraulic property at the Land Surface Model grid scale on runoff generation at two vegetated catchments.

Results from Noah-MP modeling for (a–c) the Haast river basin (NZ) and (d–f) the Auzon catchment (France). **a, d** Probability distributions of texture-based (black) and structure-corrected (red) saturated hydraulic conductivity values; the black/red dashed lines show median values for the texture/structure cases, respectively, while the blue line shows the 99th percentile of the rainfall rate (from hourly data in **b, e**). **b, e** Time series of the mean cumulative runoff (black = texture, red = structure) and the mean rainfall rate over the basins (blue). **c, f** Spatial distribution of the percentage decrease in mean annual runoff due to the soil structure correction ( $\Delta Q = (Q_{\text{text.}} - Q_{\text{str.}})/Q_{\text{text.}} * 100$ ).

more than 80% in densely vegetated areas (Fig. 4c, f). A more in-depth analysis at the scale of the single rainfall peaks shows that the role of soil structure only becomes predominant during strong rainfall events (see also Supplementary Fig. S12), as exemplified in the resulting positive correlation between rainfall intensity and runoff difference between texture- and structure-based parameterizations (Fig. 5a, b). Lastly, we observe that the parameterization proposed here recovers the positive feedback between vegetation cover and infiltration as observed in previous studies<sup>33,34</sup>. This is exemplified in Fig. 5c, d showing a higher reduction in runoff (i.e., more infiltration) due to structure corrections in areas of the catchments characterized by higher LAI values.

## Conclusions

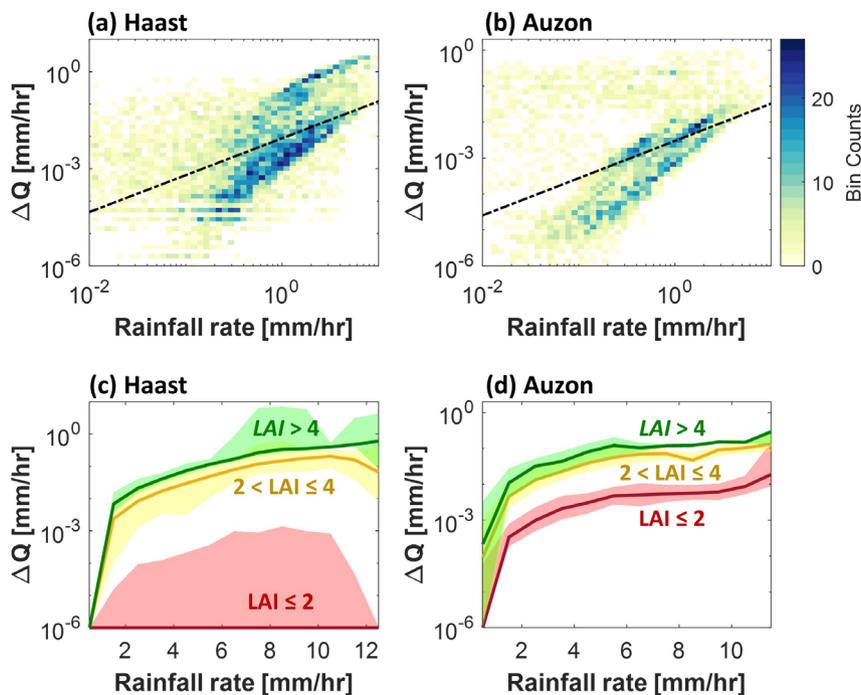
The growing need for highly resolved and reliable representation of Earth surface fluxes draws attention to the poorly constrained yet critical role of SHPs. Observations at different scales suggest that the simple paradigm of using soil texture as a main predictor of SHPs is no longer tenable and factors such as soil structure, vegetation, and scale must be integrated into the parameterization of modern LSMs<sup>15,52</sup>. Unfortunately, at this stage, we are not able to assess the potential gain in model performance that could be achieved by considering soil structure corrections, but the blueprint presented in this study offers a physically-based and systematic approach to incorporating these traits while relying, to a certain degree, on empirical relations to link vegetation and soil structure. The pragmatic approach presented here lays the foundation for future experimental and modeling efforts to include other factors that affect land surface parameterization.

We evaluated soil structure effects on landscape hydrological response focusing on rainfall partitioning to infiltration and runoff locally and, over larger heterogeneous domains, demonstrating that the upscaled parameterizations must consider spatial

correlations between vegetation and soil texture. This stands in contrast with other approaches such as flux-matching constraints<sup>23</sup> that are likely to be of limited application for correlated soil texture and vegetation, as estimated parameters would be destined to remain “associated” with a particular grid cell, thus limiting parameter transferability in space and time. The framework proposed in this study (Fig. 1) could either directly use soil and vegetation information to modify PTF-derived SHPs, or provide guidance for a parameterization that accounts for the statistical description of the sub-grid variability of the original soil parameters and vegetation attributes. Such parameterization would remain valid across scales, thus providing a direct link to current methodologies for upscaling parameter fields for distributed hydrological modeling<sup>22,24</sup>, such as the multiscale parameter regionalization<sup>22,23</sup>.

The study attempts to overcome the agricultural land bias that dominates soil survey information and remains in the core of nearly all SHPs<sup>11,12</sup>. By explicitly considering potential effects of vegetation on soil structure and SHPs and selecting parameters that are scale-informed and integrate the hydrologic response at scales of interest for modeling applications, we bridge the present gap introduced by samples from arable lands. Specifically, the study may offer a first step towards a blueprint for future spatially informed and scale-appropriate SHPs that integrate soil properties with vegetation and other landscape attributes including their sub-grid spatial organization. Within this context, we argue that the present practice of using sample values and laboratory measurements for regional and global scales applications must be revised, especially for conditions where vegetation-promoted soil structure effects are prominent and are associated with high rainfall intensities<sup>3,15,35</sup>.

Future studies should address the challenge of a more detailed quantification of functional relationships between SHPs and remotely sensed vegetation attributes (e.g., more systematic measurements of vegetation-infiltration-soil metrics) across different biomes (to better



**Fig. 5** Effect of vegetation cover and rainfall rates on runoff generation for structure and unstructured soil hydraulic property at catchment scale. Difference in runoff (as average over the catchment) for texture only and soil structure parameterization ( $\Delta Q = Q_{\text{text.}} - Q_{\text{str.}}$ ) as a function of rainfall rate over the **a** Haast and **b** Auzon catchments. The color codes the number of observations in each rectangular bin, the fitted black lines have equation  $\Delta Q = 0.009 \cdot r^{1.14}$  (with  $R^2 = 0.2$ ) and  $\Delta Q = 0.003 \cdot r^{1.04}$  (with  $R^2 = 0.3$ ) for the Haast and Auzon catchments, respectively ( $r$  being the rainfall rate). **c, d** Difference in instantaneous runoff ( $\Delta Q = Q_{\text{text.}} - Q_{\text{str.}}$ ) for grid cells with different maximum annual LAI (green =  $\text{LAI} > 4$ , yellow =  $2 < \text{LAI} \leq 4$ , red =  $\text{LAI} \leq 2$ ) for the **c** Haast and **d** Auzon basins. Solid lines show the median, while the shaded areas mark the region between the 25th and 75th percentiles. Values were computed considering peaks in rainfall (local maxima from hourly data) and the respective runoff in each grid cell—for simplicity, no time lag between rainfall peaks and runoff was considered.

constrain the parameters in Eq. (1)). In fact, while the parameterization introduced here is based on limited information, the framework is easily updatable as more data becomes available across soil types, vegetation covers, and climates. Additionally, the analysis of correlations between natural vegetation and soil properties across scales and biomes will be of interest to define the range of applicability of mean parameter values in structure corrected SHPs as well as the conditions under which generic relations can be derived for transferring parameters across scales and locations. The quantification and analysis of such feedbacks is needed for the development of effective parameterizations to be used in LSMs<sup>21,23,25</sup> accounting for the spatial heterogeneities in SHPs due to soil and vegetation distributions. Furthermore, for certain conditions and parameters, covariation and spatial organization at the sub-grid scale preclude derivation of generic (space invariant) PTFs and we may consider linking parameters with a particular grid cell (as presently done in machine learning-based digital soil mapping<sup>45,52,53</sup>).

The work here offers a framework for injecting vegetation-promoted soil structure effects in SHP parameterization, yet analyses of other functional traits (e.g., topography, spatial variability of rainfall) will be required for establishing more complete causal links between landscape attributes and heterogeneities in physical properties, as envisioned in previous studies<sup>21,25</sup>. This will provide a mechanistic strategy for model parameterization across scales, rather than ad hoc tuning and empirical calibration. Experience from other large scale studies shows that improved mechanistic representation of SHP may not directly improve model performance<sup>12,15,24</sup> due to various confounding processes and scale-appropriate parameters, yet such realism provides a path forward for updating the representation

as more information becomes available (updatability) and enables systematic attribution and adjustment of parameterization for more reliable LSMs under future climate scenarios.

## Methods

**Noah-MP LSM applications.** Simulations for the Haast and Auzon catchments were performed using the Noah-MP LSM<sup>54</sup> within HRLDAS (High-Resolution Land Data Assimilation System<sup>55</sup>), with the updated version 4.0.1 (available at <https://github.com/NCAR/hrlldas-release>)—see Supplementary Methods for additional model details. For these applications, we retrieved spatially distributed soil properties from the 1 km SoilGrids dataset<sup>56</sup>. Soil profiles were discretized in four layers (0–0.15 m, 0.15–0.6 m, 0.6–1.0 m, and 1.0–2.0 m) and properties for each layer were estimated from aggregating soil characteristics at various depths (as available from SoilGrids). Soil hydraulic properties of each layer were calculated by applying PTFs using Rosetta 3<sup>57</sup>. Soil structure corrections were introduced by modifying the saturated hydraulic conductivity based on Eqs. (1) and (2) with  $V$  equal to the maximum annual LAI (the use of maximum annual LAI in the LSM applications is motivated by the fact that in the Auzon catchment there are some deciduous forests and the use of mean annual LAI may underestimate the underground biomass). LAI values were extracted from the 1 km 8-day MODIS Leaf Area Index products<sup>58</sup>. Atmospheric forcing data, including wind speed, air temperature, air pressure, specific humidity, radiation, and precipitation were obtained from ECMWF ERA5-Land reanalysis dataset (<https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=form>), with horizontal resolution of 9 km and temporal resolution of 1 h.

## Data availability

The SoilGrids database and MODIS LAI data were retrieved from <https://soilgrids.org> and <https://earthdata.nasa.gov>, respectively. For the Noah-MP model application, atmospheric forcing data were retrieved from <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=form>. The SWIG database and the data from Weynants et al.<sup>9</sup> are available for download at <https://doi.org/10.1594/PANGAEA.885492> and <https://doi.org/10.1594/PANGAEA.879233>, respectively. Summary tables containing data used in Fig. 2b, c and Supplementary Fig. S1 are provided in the Supplementary Data 1 and are available at <https://doi.org/10.5281/zenodo.4741920>.

**Code availability**

The Noah-MP model code can be downloaded at <https://github.com/NCAR/hrldas-release>.

Received: 15 September 2020; Accepted: 11 May 2021;

Published online: 03 June 2021

**References**

- Espino, A., Mallants, D., Vanclooster, M. & Feyen, J. Cautionary notes on the use of pedotransfer functions for estimating soil hydraulic properties. *Agr. Water Manag.* **29**, 235–253 (1996).
- Vereecken, H. et al. Using pedotransfer functions to estimate the van Genuchten–Mualem soil hydraulic properties: A review. *Vadose Zone J.* **9**, 795–820 (2010).
- Gutmann, E. D. & Small, E. E. A comparison of land surface model soil hydraulic properties estimated by inverse modeling and pedotransfer functions. *Water Resour. Res.* <https://doi.org/10.1029/2006WR005135> (2007).
- De Lannoy, G. J. M., Koster, R. D., Reichle, R. H., Mahanama, S. P. P. & Liu, Q. An updated treatment of soil texture and associated hydraulic properties in a global land modeling system. *J. Adv. Model. Earth Syst.* **6**, 957–979 (2014).
- Vereecken, H. et al. Infiltration from the pedon to global grid scales: An overview and outlook for land surface modelling. *Vadose Zone J.* **18**, 1–53 (2019).
- Schaap, M. G., Leij, F. J. & Van Genuchten, M. T. Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *J. Hydrol.* **251**, 163–176 (2001).
- Wagner, B. et al. Evaluation of pedo-transfer functions for unsaturated soil hydraulic conductivity using an independent data set. *Geoderma* **102**, 275–297 (2001).
- Saxton, K. E. & Rawls, W. J. Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Sci. Soc. Am. J.* **70**, 1569–1578 (2006).
- Weynants, M., Vereecken, H. & Javaux, M. Revisiting Vereecken pedotransfer functions: Introducing a closed-form hydraulic model. *Vadose Zone J.* **8**, 86–95 (2009).
- Wösten, J. H. M., Lilly, A., Nemes, A. & Le Bas, C. Development and use of a database of hydraulic properties of European soils. *Geoderma* **90**, 169–185 (1999).
- Puhlmann, H. & von Wilpert, K. Pedotransfer functions for water retention and unsaturated hydraulic conductivity of forest soils. *J. Plant Nutr. Soil Sci.* **175**, 221–235 (2012).
- Or, D. The tyranny of small scales – on representing soil processes in global land surface models. *Water Resources Res.* <https://doi.org/10.1029/2019WR024846> (2019).
- Vereecken, H., Maes, J., Feyen, J. & Darius, P. Estimating the soil moisture retention characteristic from texture, bulk density, and carbon content. *Soil Sci.* **148**, 389–403 (1989).
- Lange, B., Lüscher, P. & Germann, P. Significance of tree roots for preferential infiltration in stagnant soils. *Hydrol. Earth Syst. Sci.* **13**, 1809–1821 (2009).
- Fatichi, S. et al. Soil structure is an important omission in Earth System Models. *Nat. Commun.* **11**, 522 (2020).
- Bonan, G. B. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science* **320**, 1444–1449 (2008).
- Klemeš, V. Conceptualization and scale in hydrology. *J. Hydrol.* **65**, 1–23 (1983).
- Dooge, J. C. I. Looking for hydrologic laws. *Water Resour. Res.* **22**, 46S–58S (1986).
- Blöschl, G. & Sivapalan, M. Scale issues in hydrological modelling: a review. *Hydrol. Process.* **9**, 251–290 (1995).
- Beven, K. The limits of splitting: Hydrology. *Sci. Total Environ.* **183**, 89–97 (1996).
- McDonnell, J. J. et al. Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. *Water Resour. Res.* <https://doi.org/10.1029/2006WR005467> (2007).
- Samaniego, L., Kumar, R. & Attinger, S. Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resour. Res.* <https://doi.org/10.1029/2008WR007327> (2010).
- Samaniego, L. et al. Toward seamless hydrologic predictions across spatial scales. *Hydrol. Earth Syst. Sci.* **21**, 4323–4346 (2017).
- Imhoff, R. O., van Verseveld, W. J., van Osnabrugge, B. & Weerts, A. H. Scaling point-scale (pedo) transfer functions to seamless large-domain parameter estimates for high-resolution distributed hydrologic modeling: An example for the Rhine river. *Water Resour. Res.* **56**, e2019WR026807 (2020).
- Sivapalan, M. Pattern, process and function: elements of a unified theory of hydrology at the catchment scale. In Anderson, M. G. (ed.) *Encyclopaedia of Hydrologic Sciences*, vol. 1, chap. 13, 193–219 (John Wiley, Hoboken, NJ, 2005).
- Thompson, S. E., Harman, C. J., Heine, P. & Katul, G. G. Vegetation–infiltration relationships across climatic and soil type gradients. *J. Geophys. Res.: Biogeosci.* <https://doi.org/10.1029/2009JG001134> (2010).
- Wilson, T. G. Effects of Vegetation and Infiltration Feedbacks on Hydrologic Partitioning and Droughts. Ph.D. thesis (2017).
- Hassler, S. K., Zimmermann, B., van Breugel, M., Hall, J. S. & Elsenbeer, H. Recovery of saturated hydraulic conductivity under secondary succession on former pasture in the humid tropics. *Forest Ecol. Manag.* **261**, 1634–1642 (2011).
- Scheffler, R., Neill, C., Krusche, A. V. & Elsenbeer, H. Soil hydraulic response to land-use change associated with the recent soybean expansion at the Amazon agricultural frontier. *Agric. Ecosyst. Environ.* **144**, 281–289 (2011).
- Niemeyer, R. J., Fremier, A. K., Heinse, R., Chávez, W. & DeClerck, F. A. J. Woody vegetation increases saturated hydraulic conductivity in dry tropical Nicaragua. *Vadose Zone J.* **13**, 1–11 (2014).
- Chandler, K. R., Stevens, C. J., Binley, A. & Keith, A. M. Influence of tree species and forest land use on soil hydraulic conductivity and implications for surface runoff generation. *Geoderma* **310**, 120–127 (2018).
- Lvovich, M. I. World water resources and their future (American Geophysical Union, 1979).
- Dunne, T., Zhang, W. & Aubry, B. F. Effects of rainfall, vegetation, and microtopography on infiltration and runoff. *Water Res. Res.* **27**, 2271–2285 (1991).
- Stone, J. J., Paige, G. B. & Hawkins, R. H. Rainfall intensity-dependent infiltration rates on rangeland rainfall simulator plots. *Trans. ASABE* **51**, 45–53 (2008).
- Van Looy, K. et al. Pedotransfer functions in Earth system science: challenges and perspectives. *Rev. Geophys.* **55**, 1199–1256 (2017).
- Tuller, M. & Or, D. Unsaturated hydraulic conductivity of structured porous media A review of liquid configuration–based models. *Vadose Zone J.* **1**, 14–37 (2002).
- Rahmati, M. et al. Development and analysis of the Soil Water Infiltration Global database. *Earth Syst. Sci. Data* **10**, 1237–1263 (2018).
- Philip, J. R. The theory of infiltration: the infiltration equation and its solution. *Soil Sci.* **83**, 345–358 (1957).
- Nielsen, D. R. J. W., Biggar, J. W. & Erh, K. T. et al. Spatial variability of field-measured soil-water properties. *Hilgardia* **42**, 215–259 (1973).
- Smith, R. E. & Dieckkrüger, B. Effective soil water characteristics and ensemble soil water profiles in heterogeneous soils. *Water Res. Res.* **32**, 1993–2002 (1996).
- Govindaraju, R. S., Corradini, C. & Morbidelli, R. A semi-analytical model of expected areal-average infiltration under spatial heterogeneity of rainfall and soil saturated hydraulic conductivity. *J. Hydrol.* **316**, 184–194 (2006).
- Craig, J. R., Liu, G. & Soulis, E. D. Runoff–infiltration partitioning using an upscaled Green–Ampt solution. *Hydrol. Proc.* **24**, 2328–2334 (2010).
- DAAC ORNL MODIS and VIIRS Land Products Global Sub-setting and Visualization Tool (2018).
- Myneni, R., Knyazikhin, Y. & Park, T. MOD15A2H MODIS/terra leaf area index/FPAR 8-day L4 global 500 m SIN grid V006. NASA EOSDIS Land Processes DAAC (2015).
- Hengl, T. et al. SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One* **12**, e0169748 (2017).
- Langhans, C. et al. Experimental rainfall–runoff data: Reconsidering the concept of infiltration capacity. *J. Hydrol.* **399**, 255–262 (2011).
- Jowitt, P. W. A conceptual systems model of rainfall–runoff on the Haast River. *J. Hydrol. (New Zealand)* **38**, 121–144 (1999).
- Kerr, T. The contribution of snowmelt to the rivers of the South Island, New Zealand. *J. Hydrol. (New Zealand)* **52**, 61–82 (2013).
- Gabrielli, C. P., Morgenstern, U., Stewart, M. K. & McDonnell, J. J. Contrasting groundwater and streamflow ages at the Maimai watershed. *Water Resour. Res.* **54**, 3937–3957 (2018).
- Ryan, M. T. et al. Vegetation and climate in Southern Hemisphere mid-latitudes since 210 ka: new insights from marine and terrestrial pollen records from New Zealand. *Quaternary Sci. Rev.* **48**, 80–98 (2012).
- Nord, G. et al. A high space-time resolution dataset linking meteorological forcing and hydro-sedimentary response in a mesoscale Mediterranean catchment (Auzon) of the Ardèche region, France. *Earth Syst. Sci. Data* **9**, 221–249 (2017).
- Gupta, S., Lehmann, P., Bonetti, S., Papritz, A. & Or, D. Global prediction of soil saturated hydraulic conductivity using random forest in a covariate-based geotransfer function (cogtf) framework. *J. Adv. Model. Earth Syst.* **13**, e2020MS002242 (2021).
- Gupta, S., Hengl, T., Lehmann, P., Bonetti, S. & Or, D. Soilscatdb: global database of soil saturated hydraulic conductivity measurements for geoscience applications. *Earth Syst. Sci. Data* **13**, 1593–1612 (2021).
- Niu, G.-Y. et al. The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *J. Geophys. Res. Atmosp.* <https://doi.org/10.1029/2010JD015139> (2011).

55. Chen, F. et al. Description and evaluation of the characteristics of the NCAR high-resolution land data assimilation system. *J. Appl. Meteorol. Clim.* **46**, 694–713 (2007).
56. Hengl, T. et al. SoilGrids1km – global soil information based on automated mapping. *PLoS One* **9**, e114788 (2014).
57. Zhang, Y. & Schaap, M. G. Weighted recalibration of the Rosetta pedotransfer model with improved estimates of hydraulic parameter distributions and summary statistics (Rosetta3). *J. Hydrol.* **547**, 39–53 (2017).
58. Yuan, H., Dai, Y., Xiao, Z., Ji, D. & Shangquan, W. Reprocessing the MODIS Leaf Area Index products for land surface and climate modelling. *Remote Sens. Environ.* **115**, 1171–1187 (2011).
59. Saxton, K. E., Rawls, W. J., Romberger, J. S. & Papendick, R. I. Estimating generalized soil-water characteristics from texture 1. *Soil Sci. Soc. Am. J.* **50**, 1031–1036 (1986).

### Acknowledgements

We acknowledge support from ETH Zürich research grant ETH-18 18-1 “The omission of soil structure in parameterization of land-surface models and the need for a new generation of pedotransfer functions”. We thank Peter Lehmann and Surya Gupta for retrieving some of the data used in Fig. 2c of the manuscript.

### Author contributions

S.B. and D.O. designed research, discussed results, and wrote the paper. S.B. performed the data analysis and modeling. Z.W. performed the Noah-MP simulations.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s43247-021-00180-0>.

**Correspondence** and requests for materials should be addressed to S.B.

**Peer review information** Primary handling editor: Joshua Dean, Joe Aslin, Clare Davis

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021