SPECIAL
Process-Oriented Model Diagnostics
COLLECTION

⌀Improving the Representation of Subsurface Water Movement in Land Models

SHERVAN GHARARI AND MARTYN P. CLARK

University of Saskatchewan Coldwater Laboratory, Canmore, Alberta, Canada

NAOKI MIZUKAMI

National Center for Atmospheric Research, Boulder, Colorado

JEFFERSON S. WONG

Global Institute for Water Security, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

ALAIN PIETRONIRO

Environment and Climate Change Canada, Saskatoon, Saskatchewan, Canada

HOWARD S. WHEATER

Global Institute for Water Security, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

ABSTRACT

Land models are increasingly used and preferred in terrestrial hydrological prediction applications. One reason for selecting land models over simpler models is that their physically based backbone enables wider application under different conditions. This study evaluates the temporal variability in streamflow simulations in land models. Specifically, we evaluate how the subsurface structure and model parameters control the partitioning of water into different flow paths and the temporal variability in streamflow. Moreover, we use a suite of model diagnostics, typically not used in the land modeling community to clarify model weaknesses and identify a path toward model improvement. Our analyses show that the typical land model structure, and their functions for moisture movement between soil layers (an approximation of Richards equation), has a distinctive signature where flashy runoff is superimposed on slow recessions. This hampers the application of land models in simulating flashier basins and headwater catchments where floods are generated. We demonstrate the added value of the preferential flow in the model simulation by including macropores in both a toy model and the Variable Infiltration Capacity model. We argue that including preferential flow in land models is essential to enable their use for multiple applications across a myriad of temporal and spatial scales.

1. Introduction

Land models have the potential to be used for a bewildering large number of applications. While land models (historically called land surface models) were initially developed to provide the lower boundary condition for atmospheric models (Manabe 1969), land models have increased in complexity and offer new opportunities to simulate the terrestrial hydrologic cycle (Clark et al. 2015; Archfield et al. 2015). Land models are attractive for hydrologic applications because they provide a more comprehensive representation of the dominant biogeophysical and biogeochemical processes controlling the water balance (Fan et al. 2019), as modern land models simulate coupled energy, water, carbon, and nitrogen cycle.

Despite the potential benefit of land models for various scientific and operational purposes, limited work has been done to evaluate the representation of hydrologic

⌀ Denotes content that is immediately available upon publication as open access.

Corresponding author: Shervan Gharari, shervan.gharari@usask.ca

processes in land models. Key challenges include the following:

1) Land models typically have a fairly simplistic subsurface structure, that is, the vertical discretization of the soil column into layers to enable application of Fourier's law and Darcy's law to describe the diffusion of heat and water through the soil column. Such discretization does not adequately represent the heterogeneity observed in nature, such as the heterogeneity in vertical/sloped flow paths.

2) The land modeling community has given limited attention to parameter estimation and uncertainty analysis. While some parameters in land models can be related to geophysical attributes, many parameters in land models are often set to spatially constant values based on limited experimental data or are based on order-of-magnitude considerations. In this approach, model parameters are often treated as physical constants, limiting the agility of process-based models (Mendoza et al. 2015; Cuntz et al. 2016). Limited work has been done to rigorously evaluate the hydrological meaning and impact of the model parameters.

3) It is challenging to evaluate land models. The current generation of land models includes a large number of processes for which there is limited evaluation data. Moreover, land models are typically run at spatial and temporal scales that differ from the scale of measurement (e.g., flux towers are not representative of larger areas; streamflow data provides limited insights on spatial variability within the catchment).

Land models also have intensive computational requirements and data needs. On the data side, land models require a myriad of geophysical data (digital terrain data, soil maps, vegetation maps, etc.), which are obtained from a mix of remote sensing and observational syntheses; the forcing data in land models includes data on precipitation, temperature, humidity, wind speed, solar and longwave radiation, and pressure, obtained from a mix of numerical weather prediction model reanalyses and station observations. Such intensive data requirements impose complicated model workflows that constrain capabilities for model analysis. On the computational side, land models are typically run over large geographical domains, at high spatial resolution, and subdaily time steps in order to simulate the diurnal variability in energy fluxes. Such intensive computational requirements limit capabilities for model analysis because it is not computationally feasible to produce a large number of model simulations.

In this paper we evaluate the temporal variability in streamflow simulations using the commonly assumed
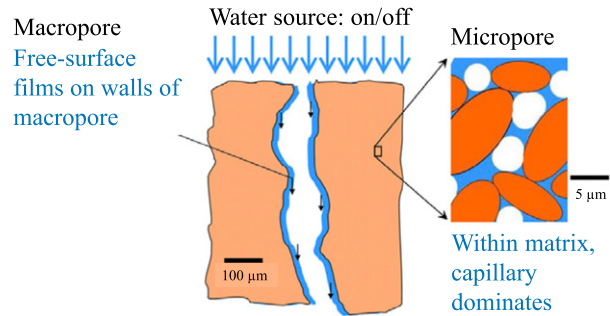

FIG. 1. Illustration of water movement in micropore and macropores. Based on the figure in Nimmo and Mitchell (2013) with the permission of the publisher.

subsurface structure of land models. We directly address the challenges defined above: we evaluate how the subsurface structure and model parameters control the partitioning of water into different flow paths and the temporal variability in streamflow, and we use a suite of model diagnostics, typically not used in the land modeling community, to clarify model weaknesses and identify a path toward model improvement.

The paper is organized as follows. In section 2 we review the modeling approaches commonly used in land models to simulate water movement in soils. In section 3 we describe our methodology, including the model descriptions and model evaluation strategy. In section 4 we present the result of each model and the interpretation of model parameters from streamflow simulations and recession analyses. The paper concludes in section 5 with discussion on the findings of this paper and possible paths forward for future model development.

2. Review of alternative approaches to simulate water movement in soils

The storage and transmission of water in soil can be described in two separate domains 1) micropores, the small pore spaces within the soil matrix, and 2) macropores, where water moves through cavities/openings developed by physical, chemical and/or biological process in soil structure. Figure 1 illustrates both the slow diffusion of water in micropores (e.g., the soil matrix where the pore space is smaller than $5\,\mu m$), as well as the fast downward flow of water through macropores (e.g., the soil cavities of around $100\,\mu m$). The flow of water through macropores is difficult to quantify because it depends on a myriad of factors such as biological processes, tree density and type, density of earthworms, bedrock topography, human construction of artificial drainage, and so forth. The heterogeneity of macropore flow and the lack of information on macropore structure and connectivity make it difficult to simulate macropore

flow directly (Beven and Germann 1982; Koide and Wheater 1992; Wheater et al. 1993; Mirus et al. 2011; Mirus and Nimmo 2013; Harman et al. 2009; Davies et al. 2013).

One of the major differences between bucket-style rainfall–runoff models, commonly referred to as conceptual rainfall–runoff models, and land models, is the representation of the fast and slow flow, conceptualized based on the macropore/micropore water movement. In commonly used hydrological models such as HBV, FLEX, and HYMOD, the precipitation is first split into two pathways: 1) the water that is stored, absorbed, and stays within the micropore matrix and 2) the water that is transmitted through the soil matrix via macropores. The water transmitted through macropores can have multiple flow pathways, for example, a "faster" flow pathway where flow through macropores is delivered directly to the streams, and a "slower" flow pathway where macropore flow feeds a baseflow reservoir. Some model structures also include the percolation, or slow movement of water, from the micropore matrix to the baseflow reservoir. These simpler models are often presented as a collection of buckets, and the physical meaning of the model structure (e.g., the micropore matrix, macropore flow) is not explicit.

In land models the precipitation is typically partitioned into two flow pathways: the surface runoff that is routed to the river network, and the infiltration into the soil (i.e., infiltration into the micropore matrix). Modern land models do not typically represent macropore flow at all (Clark et al. 2015). Most land models have a similar treatment of the storage and transmission of water through micropores (Boone and Wetzel 1996; Lee and Abriola 1999), that is, most land models use Richards' equation to describe the gravity and capillary fluxes in the micropore matrix, but differ in the constitutive functions that are used to relate matric head to volumetric liquid water content and hydraulic conductivity (van Genuchten, Brooks and Corey, etc.) and the aspects of the numerical solution (e.g., the vertical discretization of the soil column, methods to interpolate fluxes to layer interfaces, and the time stepping scheme). The drainage of water from the bottom of the micropore matrix is the slow component of flow, which may be used to recharge a conceptual aquifer. Modeling the heterogeneity of flow through micropores and macropores is a key unmet challenge for the land modeling community (Clark et al. 2015; Rahman and Rosolem 2017).

More generally, model simulations of the storage and transmission of water through micropores and macropores can be classified into three discrete modeling approaches, all of which are variants of the complete model in Fig. 1.

1) The approach typically used in land models where the storage and transmission of water in the unsaturated zone is only based on diffusion through micropores. The model neglects any representation of macropores.
2) In the approach where flow only occurs in macropores, the water that infiltrates into the micropore matrix stays in the system and can be depleted only by transpiration. This approach is mostly used in catchment scale bucket-style rainfall–runoff models.
3) The approach that simulates both micropore and macropore flow.

3. Methods

In this paper we use a three-tiered modeling strategy to evaluate how variations in model structure and model parameters affect the temporal variability of surface runoff and baseflow. First, we use a toy model to clearly illustrate how inclusion of macropores can increase the temporal variability of streamflow. Next, we focus on variations of the Variable Infiltration Capacity (VIC) model that include and exclude macropores. Finally, we investigate the effect of spatiotemporal model aggregation over longer time steps as well as the VIC model behavior over various climates for the catchments of continental United States based on the Catchment Attributes and Meteorology for Large-Sample Studies (CAMELS) dataset (Addor et al. 2017).

For the analyses of models' simulations we rely on the mean and standard deviation of the simulated fast and slow reacting responses. We also carry out recession analysis on the simulated flow to compare the characteristics of the recession simulation by various models.

a. Models

1) TOY MODEL

To start simple, we use toy models by following the example of work by Koster and Milly (1997) and Koster and Mahanama (2012) to isolate individual processes, namely the macropore/micropore water movement in soil. Using the toy model structure, we define three models that emphasize different contributions of macropore and micropore flow (refer to appendix A for the full description of the models). The models are as follows:

- Model$_{micro}$: This model only simulates micropore flow and lacks proper representation of macropore water movement in the soil matrix (only uniform flow through micropore flow with no macropore flow). This model is

TABLE 1. The parameter ranges of the toy model. To replicate Model$_{\text{micro}}$, D is set to 0 and for Model$_{\text{macro}}$, k_{micro} is set to 0.

Parameter	Description	Unit	Range
T_{snow}	The temperature below which precipitation is considered to be snow	°C	−2.00 to 2.00
c_{ddf}	Degree-day factor for snowmelt for a unit of temperature above T_{snow}	mm °C^{-1} day^{-1}	1.00–5.00
$S_{u\text{max}}$	Maximum soil moisture capacity	mm	50.00–500.00
f_{thr}	Fraction of maximum soil moisture below which all the snowmelt and rainfall enter the unsaturated zone.	—	0.10–0.90
f_{crit}	Fraction of maximum soil moisture below which the transpiration is limited by soil moisture	—	0.10–0.90
D	Macropore volume fraction for partitioning the runoff to the fast and slow reservoirs	—	0.00–1.00
k_{micro}	Maximum percolation rate when storage is at its maximum, $S_{u\text{max}}$	mm day^{-1}	0.00–10.00
k_{fast}	The fast reservoir coefficient	day^{-1}	0.10–0.90
k_{slow}	The slow reservoir coefficient	day^{-1}	0.01–0.10

constructed by setting the fraction of macropore flow, parameter D, to 0.

- Model$_{\text{dual}}$: This model simulates both macropore and micropore water movement.
- Model$_{\text{macro}}$: This model simulates only macropore water movement. The infiltration into the micropore matrix is only depleted by transpiration, and no percolation to more saturated soil layers (immobile water in micropore space; only macropore flow). This model is constructed by setting the maximum percolation rate k_{micro}, to 0.

For the model simulations, the parameter ranges are set as wide as possible to have as large model ensemble as possible (Table 1). Each of the variations of toy model is simulated 10 000 times from the feasible parameter space with marginal uniform distribution for every parameter.

2) LAND MODEL, VIC MODEL

The VIC model is a relatively simple land model (Wood et al. 1992; Liang et al. 1994). The process conceptualization is based on the work of Zhao et al. (1980) that computes the infiltration rate based on the saturation or storage of the catchment. For this study, we utilize VIC version 5 (Hamman et al. 2018), which is forced by seven climatic variables, namely, precipitation, pressure, wind, vapor pressure, temperature, and longwave and shortwave radiation. The soil parameters include saturated hydraulic conductivity, soil and bulk density, residual moisture, the slope of water retention curve b_{inf}, the variable infiltration capacity shape function, and soil layer depth identifying the active hydrological depth. Many of these parameters are difficult to measure, especially at larger spatial scale, and are often subjected to calibration given the observed streamflow. Apart from these parameters, and due to the fact that the VIC model is a mesoscale model implemented at large

scales, some of the soil physical parameters such as saturated hydraulic conductivity K_{sat} and the slope of water retention curve E_{exp} are also subjected to calibration to find the most "*effective*" parameter values at the scale of modeling. The VIC model also requires vegetation parameters which include stomatal resistance and leaf area index (LAI) that are used in Jarvis formulation to simulate transpiration. The Jarvis formulation considers factors such as temperature, humidity, soil moisture and solar radiation availability. The root zone depth is also indicated in the vegetation parameters.

The effective precipitation, the amount of rainfall or snowmelt that is not intercepted by vegetation foliage, either enters the micropore matrix or is partitioned into fast reacting surface runoff. The micropore water content is then transmitted between the top two soil layers using an approximation of Richards' equation (uniform across grid as the VIC model only accepts one soil type per layer per grid). The bottom (third) soil layer is considered to be the baseflow or saturated layer that generates the baseflow based on a maximum possible flux per day $D_{s\text{max}}$. Baseflow also depends on other baseflow parameters, D_s, W_s, and c. More importantly, baseflow depends on d_3 that affects the saturation level for a unit depth of soil, and therefore changes the baseflow significantly. In this study and for the sake of simplicity, we ensure that the root depths are equal to the first two top layers, avoiding the penetration of roots to the third layer that is responsible for baseflow generation.

We designed five different variations of the VIC model, in the following order:

- VIC$_{\text{original}}$: this model has the exact structure of the original VIC model. We perturb the parameters b_{inf}, k_{sat}, d_1, d_2, and the four baseflow parameters of ARNO formulations, $D_{s\text{max}}$, D_s and W_s as well as d_3, to evaluate the performance of this model behavior

TABLE 2. The range of the model parameters for various VIC models.

Parameter	Description	Unit	Range
b_{inf}	The variable infiltration curve coefficient	—	0.01–5.00
K_{sat}	Saturated hydraulic conductivity	mm day^{-1}	50.00–10 000.00
E_{exp}	The slope of water retention curve	—	3.00–12.00
d_1	The depth of the first soil layer	m	0.05–0.30
d_2	The depth of the second soil layer	m	0.30–2.00
d_3	The depth of the bottom (third) soil layer	m	0.30–2.00
D_{smax}	Maximum baseflow rate when the third layer is fully saturated	mm day^{-1}	0.00–10.00
W_s	The fraction of maximum baseflow at the point the baseflow becomes nonlinear	—	0.00–1.00
D_s	The fraction of the saturation after which the baseflow becomes nonlinear	—	0.00–1.00
D	Macropore volume fraction for partitioning the surface runoff into effective runoff and macropore water movement.	—	0.00–1.00
k_{slow}	The slow reservoir coefficient	day^{-1}	0.01–0.10

(for the sake of simplicity the power for the nonlinear behavior of the baseflow c_{exp}, is set to 2).

- VIC$_{original\_no\_baseflow}$: this model is designed to evaluate the effect of the original VIC baseflow on the streamflow simulation. In this model we assume that no baseflow module is present, and the baseflow is considered to be the moisture transmitted from the second layer. As the root depth is set to be equal to the first two top layers, the omission of the baseflow layer does not affect partitioning of precipitation to runoff, baseflow, and transpiration.

We construct three different VIC models with simplified assumption on its baseflow component. The original VIC baseflow, or the bottom layer, and its formulation is replaced by a slow linear reservoir and its coefficient (k_{slow}). These models are as follow:

- VIC$_{micro}$: this model is the closest to the original VIC model, but with the difference that the baseflow layer is replaced with a linear reservoir. The water that feeds the baseflow reservoir is transmitted through the soil layers only based on the micropore water movement assumption as similar to the original VIC model.

- VIC$_{dual}$: the other contributing flux to the slow linear reservoir is fraction of generated surface runoff that is transmitted through macropores to the slow linear reservoir. This fraction is computed using a simple macropore volume fraction D that varies between 0 and 1. The fraction of surface runoff that does not enter the baseflow reservoir directly contributes to the river network as surface runoff.

- VIC$_{macro}$: this model is similar to VIC$_{dual}$, with the difference that no micropore water movement is allowed to percolate to the bottom soil layer (immobile water in micropore space). The only flux that contributes to the saturated linear flux is macropore water movement as fraction of surface runoff, which is identified by macropore volume fraction D.

Each of the VIC variations are run with 10 000 parameter sets sampled from the feasible parameter space based on the parameter range reported in Table 2 with uniform distribution along each parameter. Before analyzing the model simulations, and to ensure that the model simulations are realistic, a constraint is applied to baseflow simulations of VIC$_{original}$ to ensure that the bottom layer is not accumulating water throughout the time period of the model simulation, which is typically the case for the original VIC baseflow formulation. Moreover, we ensure that the baseflow is not fixed at a certain value for an extended period of time which means the bottom layer might be fully saturated. We would like to avoid this condition as the saturated bottom layer, based on the VIC assumptions, passes the excess water to the upper layers, which then changes the composition of the runoff and baseflow separation. To make sure that the model states and fluxes are initialized properly, we also allow 10 years of spinup period for all the variations of VIC model.

b. Model simulations

1) SINGLE POINT SIMULATIONS

We initially evaluate single point simulations for a location in the Canadian Rockies to provide an example of model behavior. The forcing data is from WATCH-Forcing-Data-ERA-Interim (WFDEI; Weedon et al. 2014), from a grid cell centered at 48.75°N and 113.75°W. The land cover of the point simulation is considered to be forest. We use 37 years of the data, 1979–2016, from which 10 years are used for spinup to make sure that the original VIC (VIC$_{original}$) baseflow is stable as mentioned earlier, meaning that no water is accumulated over time in the baseflow layer. For the toy model we use the same dataset, WFDEI for the Canadian Rockies, but only precipitation and temperature are used as model forcings. Daily temperature and the

latitude of the region of study are used to calculate the potential evaporation based on the Hamon equation (Hamon 1960).

2) SPATIOTEMPORAL AGGREGATION OF SINGLE POINT SIMULATIONS

We perform additional simulations to evaluate the extent to which results may be masked by spatial routing and temporal aggregation.

(i) Spatial routing

Often the surface runoff and baseflow simulations from a land model are routed over a larger area using a routing scheme that aggregates the model simulation from many points into one or few point(s). It is hence important to evaluate the extent to which level this aggregation smooths or masks the model simulation at the grid level. To investigate this effect, we set up a synthetic routing model based on diffusive wave formulation and assumptions on the relation of river length and river basin size that are based on Hack's law (refer to appendix B for the detailed explanation of the routing model). We then route the fast and slow simulations of VIC$_{original}$ for basins of various sizes of 1000, 10 000, and 100 000 km^2 assuming a similar grid of 100 km^2 with no heterogeneity (parameters and forcing are the same for all the grids). We then compare the routed results for various spatial scales.

(ii) Temporal aggregation

Despite the fact that land models use simulation time steps shorter than a day, these models are often evaluated at temporal resolution of day(s), and for larger basins with human interference even longer, month(s). Similar to spatial routing, the structural deficiencies of land models may be masked by temporal aggregation. To check this effect, we simply aggregated the model output from the VIC$_{original}$ to a monthly scale and evaluate the summary metrics of the fast and slow flow.

3) CONUS SIMULATIONS

To investigate the wider behavior of the VIC model fast and slow flow components across larger spatial domain and with different forcing data, we make use of the calibrated VIC setup for the 531 individual basins in the CAMELS dataset [for the details of forcing, soil, and land cover data, please refer to Mizukami et al. (2017)]. We look into the best performing simulated runoff and baseflow for every individual basin of interest. This test is used to check if the overall VIC behavior from the single-point simulations is representative of behavior for small to mesoscale catchments across the contiguous United States that span over various climate, soil and land cover

and topographical conditions. To compare the simulated and observed hydrographs in recession behavior, we investigate the recession parameters for observed and simulated hydrographs for each individual basins.

c. Diagnostic methods and metrics

1) MEAN AND STANDARD DEVIATION

To compare the behavior of the fast and slow flow for every model simulations, commonly used summary metrics, the mean and standard deviation of the simulated time series are used. We also compare the mean of the fast and slow flow ensembles of each model visually.

2) RECESSION ANALYSIS

As a diagnostic tool, we carry out recession analysis on the simulated hydrographs. Recession analysis is one of the simplest methods to get insights on the catchment/basin characteristics based on the observed hydrograph. There have been numerous methods for the recession analysis and baseflow separation. In this study we use the widely used empirical relationship between the streamflow value and its rate of change as a power function:

$$-\frac{dQ}{dt} = aQ^b, \tag{1}$$

where a is the intercept of the relationship when Q is equal to 1 and b is the slope of the log–log plot for dQ/dt versus Q. During the recession period when there is no precipitation input, the storage–discharge relationship for the reservoir that discharges slow flow component can be expressed as

$$\frac{dS}{dt} = -Q. \tag{2}$$

Discharge Q during the recession is generally described as a function of storage in the models:

$$Q = \frac{1}{c}S^d. \tag{3}$$

The parameter c is usually referred to as a retention time of the reservoir, and c^{-1} is referred to as a reservoir coefficient. The parameters a, b, c, and d are interrelated. The relationship between the parameters in Eqs. (1) and (3) can be found by inserting Eq. (3) into Eq. (2) as follows (Clark et al. 2009):

$$a = c^{-1/d}d, \tag{4}$$
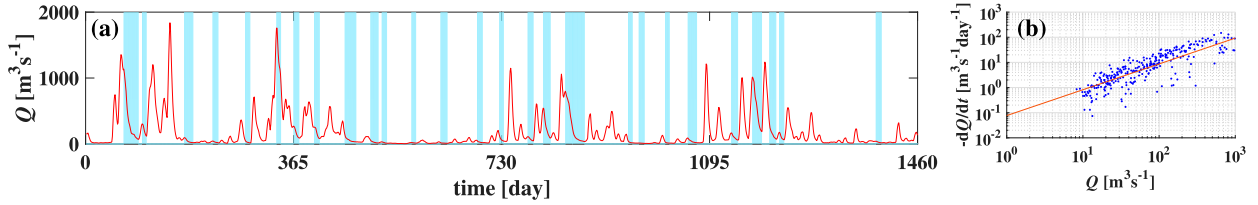
$$b = 2 - \frac{1}{d}, \tag{5}$$

FIG. 2. (a) The shaded areas are the parts of the hydrograph that are selected as recession analysis. (b) The relationship between dQ/dt and Q and its linear relationship between them with a and b equal 0.076 and 1.027, respectively. Streamflow data from Maumee River at Waterville, OH, USGS 04193500 (USGS 2019).

$$c = [a(2 - b)]^{1/(b-2)}, \quad \text{and} \tag{6}$$

$$d = \frac{1}{2 - b}. \tag{7}$$

To identify the recessions, we use a similar approach proposed by Vogel and Kroll (1992). First, we applied a moving average of five days and considered the recession part of smoothed hydrograph when the change of two consecutive time steps are less than a certain ratio (in this case less than 20% of previous time step). We retain recessions that are longer than six time steps (i.e., 6 days) for this study.

If we logarithm transform both sides of Eq. (1),

$$\log\left(-\frac{dQ}{dt}\right) = \log(aQ^b), \tag{8}$$

using the logarithm rules, we can rewrite the above equation as

$$\log\left(-\frac{dQ}{dt}\right) = \log(a) + b\log(Q), \tag{9}$$

which has the form of a linear relationship between $\log(dQ/dt)$ and $\log(Q)$ as variable from which a and b can be calculated. The parameters a and b can be calculated for every model simulation of the ensembles resulting in an ensemble of recession parameters which can shed light on model behavior. Figure 2 illustrates the graphical implementation of the recession analyses on an observed hydrograph.

4. Results

a. Toy model

Figure 3 illustrates the mean and standard deviation of the fast and slow reacting components, macropores and micropores, respectively, for each of the 10 000 model parameter sets for the three variations of the toy model. Model$_{\mathrm{micro}}$, which represents structure similar to land models, has a very distinct behavior where there is strong separation between fast and slow reacting components. The summary

metrics of the two flow pathways do not overlap in Fig. 3a which results in a two distinctive segregated fast and slow response from the model. This can be visually inspected in Fig. 3b, which shows the mean values of the baseflow and fast flow of the ensemble simulations.

Figure 3e shows the summary metrics of Model$_{\mathrm{macro}}$ that only allows for preferential flow. This model generates summary metrics of fast and slow reaction which overlap, meaning there are ensemble members with smoother transition between the model fast and slow simulations. From Fig. 3f it is clear that the slow reacting reservoir of Model$_{\mathrm{macro}}$ is able to produce baseflow fluxes with higher variability and shorter recession in comparison with the other two models.

Model$_{\mathrm{dual}}$ is an intermediate model between Model$_{\mathrm{macro}}$ and Model$_{\mathrm{micro}}$. Similar to Model$_{\mathrm{micro}}$, Model$_{\mathrm{dual}}$ also generates a slow responding baseflow with higher values compared to Model$_{\mathrm{micro}}$. The reason for this inflation in baseflow in Model$_{\mathrm{dual}}$ is that the unsaturated zone is depleted by micropore water movement, which results in lower soil moisture and eventually results in more precipitation to enter micropores and a higher contribution of baseflow in the model simulations. A closer look at the model shows that the Model$_{\mathrm{macro}}$ is able to generate peaks flows during the dry period. By contrast, Model$_{\mathrm{micro}}$ and Model$_{\mathrm{dual}}$ have lower peak flows because all or more of the effective rainfall is transmitted through the soil matrix and contributes to the flow as a diffused signal. This may hamper the power of the model for flood generation during dry periods characterized by rapid preferential flow.

Recession analyses are performed for each of the 10 000 simulations. Figure 4 shows the parameters a and b for the simulations for each of the three models. The result shows that Model$_{\mathrm{micro}}$ and Model$_{\mathrm{dual}}$ are inherently different in their recession behavior. Model$_{\mathrm{micro}}$, which is similar to the land models and emphasize micropore movement of water, generate recessions with parameter a less than 0.05 day^{-1} (or parameter c longer than 20 days). Its center of mass, mean, for parameter a is around 0.02, which means recession of
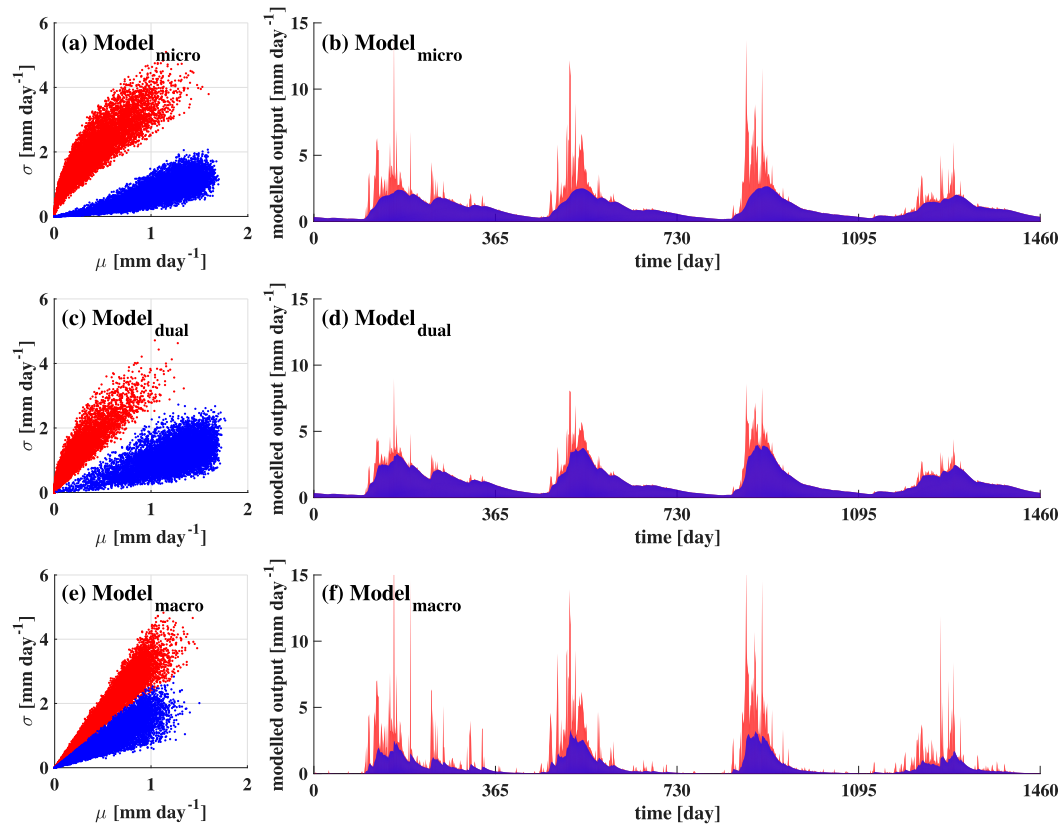
FIG. 3. The summary metrics, mean and standard deviation, for fast (red) and slow (blue) reacting components of the simulated hydrographs by model (a) Model$_{micro}$, (c) Model$_{dual}$, and (e) Model$_{macro}$ and the mean of the fast and slow simulations ensemble over time by (b) Model$_{micro}$, (d) Model$_{dual}$, and (f) Model$_{macro}$.

approximately 50 days and longer. Therefore the structure of Model$_{micro}$ constrains model capabilities to simulate flashier catchments with retention time less than 20 days. On the other hand, the model with macropore flow, or preferential flow, similar to widely used catchment scale bucket-style rainfall–runoff models, are able to produce the retention parameter of a up to $0.2 \, \text{day}^{-1}$ or c values as low as 5 days.

Parameter b remain always above 1 as the recession is combination of a linear behaving reservoir, slow/saturated reacting reservoir and a fast component, direct contribution to streamflow. If the slow recession is also
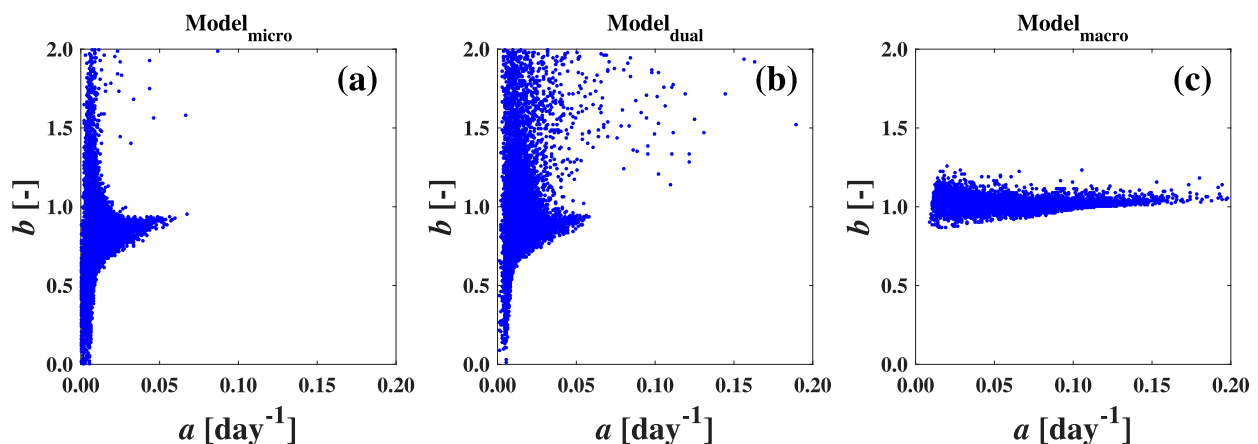

FIG. 4. Comparison of the recession analysis, a and b parameters, for the three different model.

considered to be nonlinear then the model will be able to even generate wider combination of parameter a and b. Model$_{dual}$ shows an intermediate performance of the two but tend to emphasis on longer recessions compare to Model$_{micro}$ as it was also the case in Fig. 3d.

b. Application to a land model, VIC

1) STRUCTURAL DIFFERENCES

Similar to the analyses on the toy models, the mean and standard deviation of the five variations of the VIC model are illustrated in Fig. 5. Figure 5a shows the mean and standard deviation of fast and slow flows for VIC$_{original}$ setup. The fast flow component illustrated in red are clustered around a line. This line can be represented by intercept of 0 and slope of standard deviation over mean of precipitation over the entire period of modeling. Similar to the toy model, the summary metrics of the fast and slow flows do not overlap for VIC$_{original}$. VIC$_{original\_no\_baseflow}$ is a variation of VIC$_{original}$ in which the output from the second layer is treated as baseflow. The result shows that the baseflow response becomes very flashy and even can dominate the fast-reacting component. Therefore, we add a simple linear reservoir instead of the more complex VIC$_{original}$ baseflow to account for the baseflow separation.

VIC$_{micro}$, which is similar to the VIC$_{original}$ with the difference of a simplified baseflow, acts similarly to VIC$_{original}$. The slow flow component, or baseflow, is even less diverse considering the cloud of mean and standard deviation in comparison to VIC$_{original}$ as the simplified baseflow has less degree of freedom in comparison to the VIC$_{original}$ baseflow formulations. VIC$_{dual}$, which allows both macropores and micropores water movement, has a less strong runoff reaction as more water is moved to the baseflow and therefore less soil moisture result in less runoff generation in general. This is similar to Model$_{dual}$. VIC$_{macro}$, which allows for macropore water movement only, result in higher and more diverse variability of the fast flow. This is due to the fact that the soil moisture is immobile and not depleted by the micropore movement to the baseflow and therefore the top soil layers remain wetter which result in more runoff generation compare to the VIC$_{micro}$ and VIC$_{dual}$. The macropore volume fraction D that redirects part of surface runoff to more saturated reservoir also diversifies the fast flow response of the VIC$_{macro}$. Moreover, the slow reacting component of model VIC$_{macro}$ has shorter recessions than the other model that allows for micropore water movement. VIC$_{macro}$ also generated higher peaks, which are essential in model capabilities to capture peaks during flood event especially in drier periods.

(i) Parameter effect on the spread of runoff and baseflow for VIC$_{original}$

To evaluate the effect of the original VIC model, VIC$_{original}$, parameters on the spread of mean and standard deviation presented in Figure 5a, we evaluate the change in variation in the mean-standard deviation plot by changing the target parameters. Please note that this is only an illustrative example and is not a comprehensive sensitivity analysis. The patchiness in the figures is clusters of sets of parameters that are selected for visualization in 2D space of that target parameter while two or three other parameters are varied in their ranges. Moreover, the parameters sets are not identical for each and every panel for the visualization purposes.

For the analysis, we first perturb the parameters b_{inf}, K_{sat}, d_1, and d_2, which are responsible for baseflow/runoff separation in the original VIC model, VIC$_{original}$. Increased b_{inf} will result in higher mean and variability in runoff while it reduces the mean and standard deviation of the baseflow as it redirects more water to runoff (Figs. 6a,f). Saturated hydraulic conductivity K_{sat} acts the opposite direction, increased K_{sat} results in more water to be transported to the baseflow layer, less soil moisture and therefore more infiltration to the soil and therefore reduced runoff (Figs. 6b,g). Parameters d_1 and d_2 act similarly: increasing d_1 and d_2 will decrease the mean and variation of baseflow (Figs. 6c,d,h,i). Increased depth will result in more water to be stored in the system, which in turn results in less mean and variability of the output both as runoff and baseflow. Figures 6e and 6j show the general effect of the parameters perturbation on generation of runoff and baseflow mean and variability.

Second, we look into the baseflow parameters, namely, D_s, D_{smax}, W_s, and d_3 (c_{exp} is fixed at 2). In this case, the higher D_s result in higher baseflow variation. Increased D_{smax} results in increased variation in the baseflow behavior as well. Increased W_s results in increased baseflow variation, while increased depth d_3 results in a decreased variation. The trend for D_{smax} and d_3 can be inferred directly from the VIC baseflow formulation, the sign for D_{smax} is positive and also increased depth d_3 results in lower saturation level which in turn decreases the baseflow variability. Parameters D_s and W_s, however, may show more complex behavior than described here. If the constraints were not imposed on baseflow behavior, the D_s, D_w interplay would result in more complex behavior.

(ii) Parameter effect on the spread of runoff and baseflow for VIC$_{macro}$

We carry out similar illustrative sensitivity analyses to understand the effect of each parameter on the mean
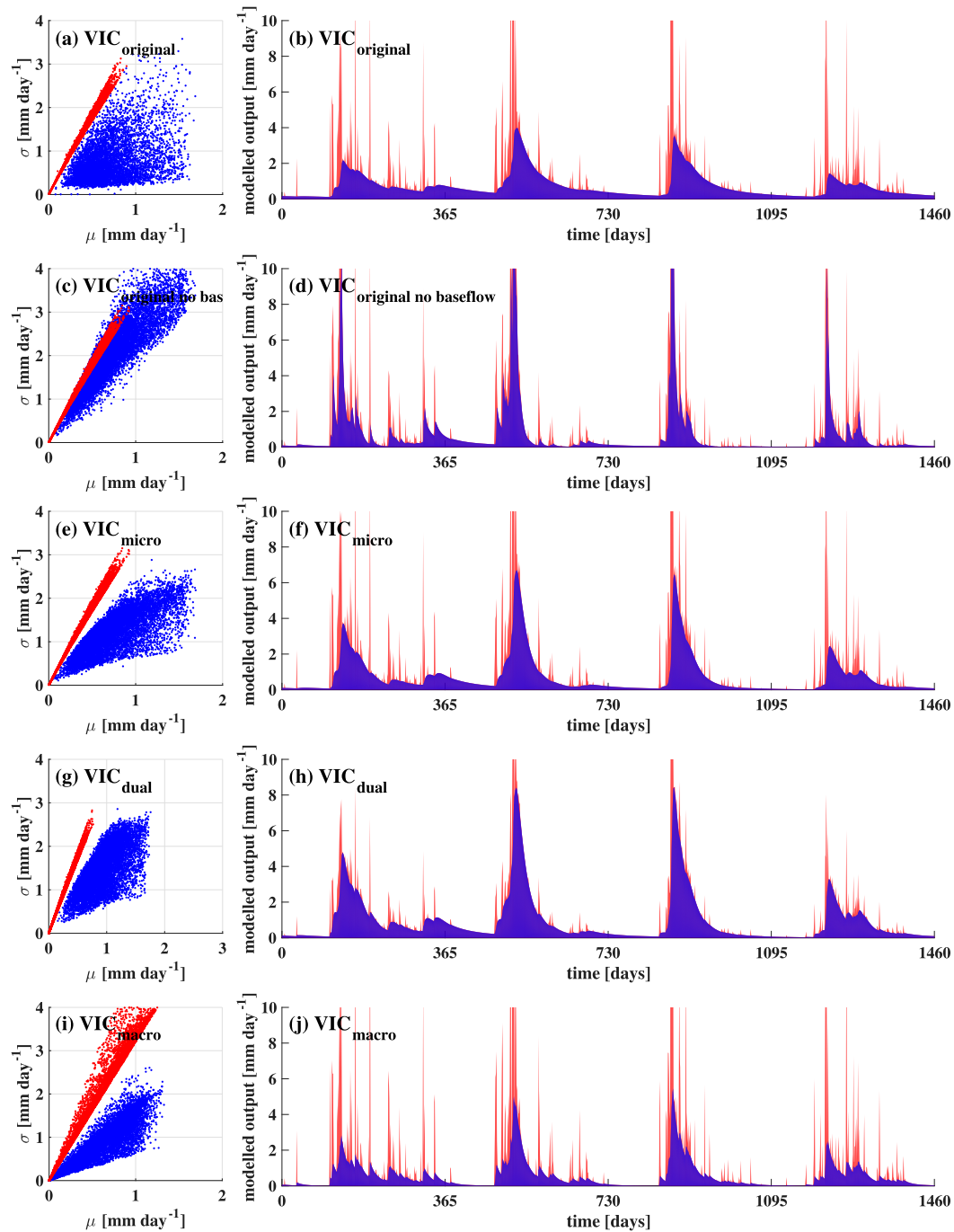
FIG. 5. The summary metrics, mean and standard deviation, for fast (red) and slow (blue) reacting components of the simulated hydrographs by model (a) VIC$_{original}$, (c) VIC$_{original\_no\_baseflow}$, (e) VIC$_{micro}$, (g) VIC$_{dual}$, and (i) VIC$_{marco}$ and simulation of the mean of the fast and slow ensemble over time by (b) VIC$_{original}$, (d) VIC$_{original\_no\_baseflow}$, (f) VIC$_{micro}$, (h) VIC$_{dual}$, and (j) VIC$_{marco}$.

and standard deviation of the fast and slow response of the VIC$_{macro}$ presented in Fig. 5i for the target parameters b_{inf}, d_2, K_{slow}, and D. For this analysis, to simplify the case, we fix d_1 at 10 cm. Increased b_{inf} results in lower variability of runoff. Higher b_{inf} results in less

water entering the soil layer, which in turn results in a dryer condition, which in turn results in less flashy behavior of the system during a rainfall event (Fig. 7a). For the baseflow, as macropore water movement is considered to be a fraction of surface runoff that is defined by
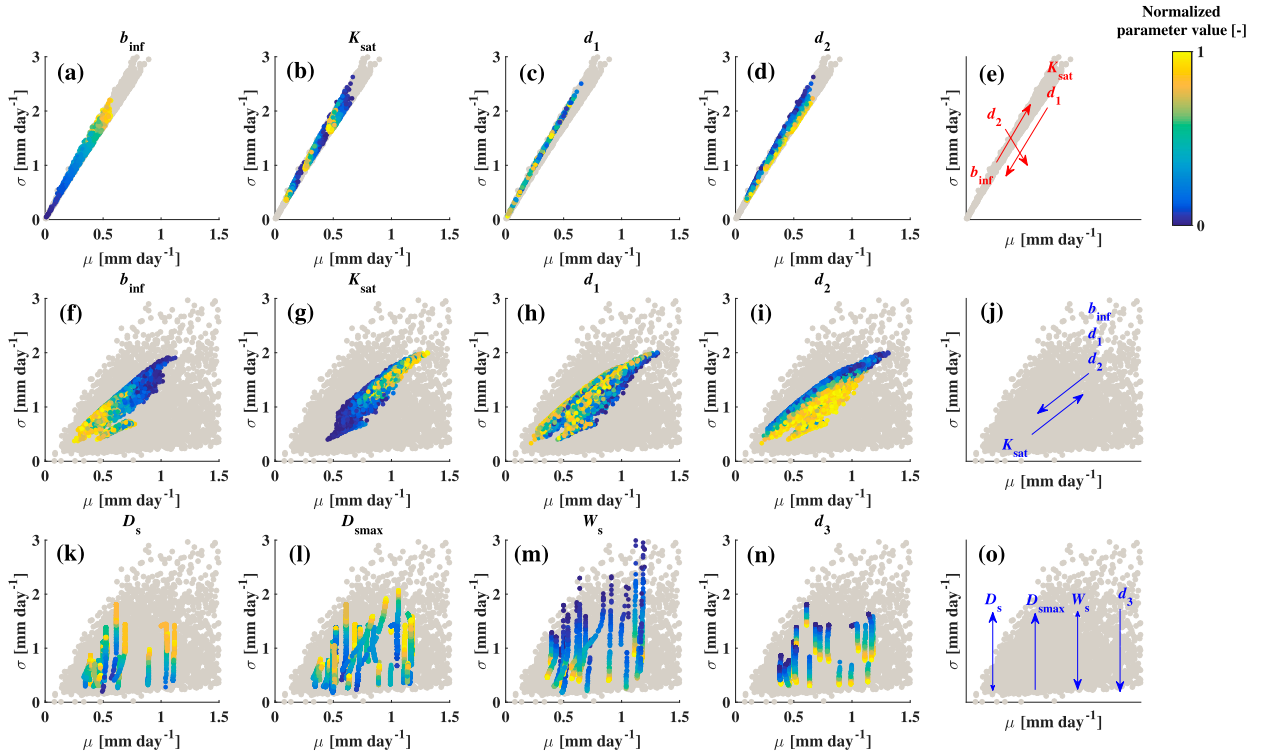
FIG. 6. The change in mean and standard deviation for the runoff (a) b_{inf}, (b) K_{sat}, (c) d_1, and (d) d_2, and the change in mean and standard deviation for the baseflow for (f) b_{inf}, (g) K_{sat}, (h) d_1, (i) d_2, (k) D_s, (l) D_{smax}, (m) W_s, and (n) d_3 for VIC$_{original}$. The overall changes in mean and standard deviation for target parameters are presented in (e), (j), and (o). The lighter color for the points means a higher value of normalized parameter ranges. Please note that this is only an illustrative example and is not a comprehensive sensitivity analyses. The patchiness in the figures is the cluster of sets of parameters that are selected for visualization in 2D space of that target parameter while two or three other parameters are varied in their ranges. Moreover, the parameters sets are not identical for each and every panel for visualization purposes.

D, increased b_{inf} results in increased mean and variability of baseflow response (Fig. 7f). Larger soil depth, d_2, translates into more storage, which results in less baseflow and runoff mean and variability. Increased K_{slow}, the simplified baseflow coefficient, results in increased baseflow variability. Increased macropore volume fraction parameter D results in more water going to baseflow and a higher mean and variability for this flux, while it results in less mean and variability for the fast flow component.

2) RECESSION BEHAVIOR

A similar recession analysis depicts that original VIC model, VIC$_{original}$, is not able to generate recessions that have an a parameter of more 0.025 than or c parameter less than 40 days, while VIC$_{macro}$ can generate recession parameter a that is as high as 0.15 or c parameter of approximately 8 days, indicating that VIC$_{macro}$ is capable to replicate the behavior of flashier basins or headwater catchments (Fig. 8). The other models overall produce lower values of parameters a than VIC$_{macro}$ even though they can produce higher value than VIC$_{original}$. This indicates that micropore

water process will put a heavy weight on generating longer recessions.

c. On the effect of routing on the model simulation

As explained earlier, spatial routing might mask the representation of subsurface processes in land models. To address this issue, we route the original single-point VIC simulation, VIC$_{original}$ through a simple routing model explained in appendix B. Figure 9 illustrates the mean and standard deviation of the routed fast and slow flow at the basin outlet for basin sizes of 1000, 10 000, and 100 000 km^2. The results indicate that the larger the basin gets, the standard deviation of fast response summary metrics are reduced and overlap with slow response summary metrics. Figure 9 also illustrates that as the basin size gets larger the hydrograph become smoother, and therefore, calibrated model for large river basin can generate a smooth looking hydrograph with a good transition between fast and slow flow while its component at every grid may not follow the same smooth transition. Since streamflow for larger basins is highly affected by the routing, the subsurface structure
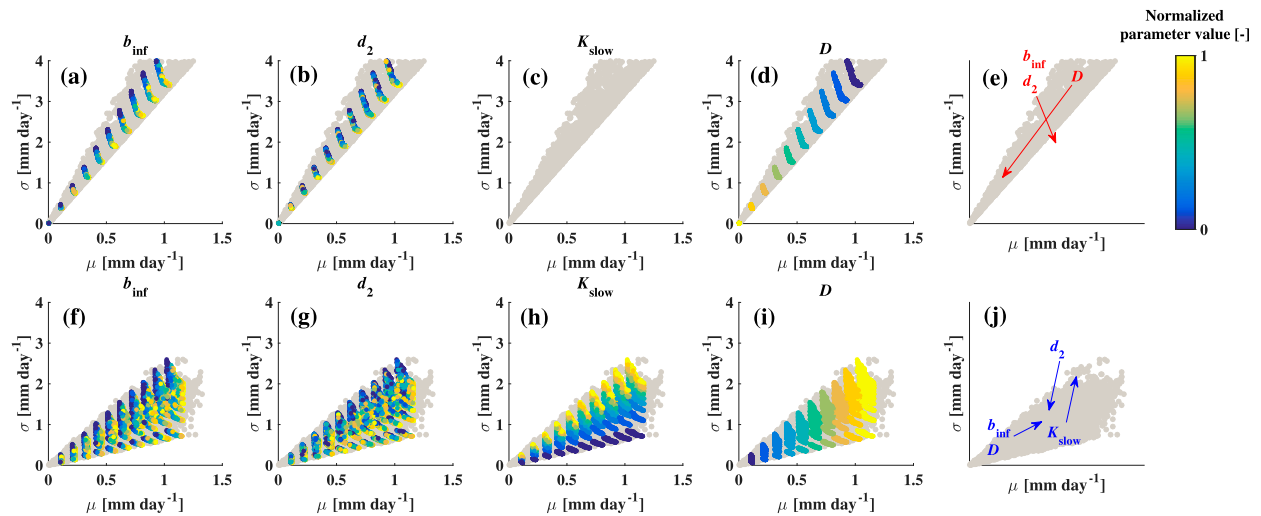
FIG. 7. The change in mean and standard deviation for the runoff (a) b_{inf}, (b) d_2, (c) K_{slow}, and (d) D and the change in mean and standard deviation for the baseflow for (f) b_{inf}, (g) d_2, (h) K_{slow}, and (i) D for VIC$_{macro}$. The overall changes in mean and standard deviation for target parameters are presented in (e) and (j). The lighter color for the points means a higher value of normalized parameter ranges. Please note that this is only an illustrative example and is not a comprehensive sensitivity analyses. The patchiness in the figures is the cluster of sets of parameters that are selected for visualization in 2D space of that target parameter while two or three other parameters are varied in their ranges.

of the model is masked by the routing formulation and river network shape. This should be taken into consideration when evaluating land model performance of spatially aggregated point simulations.

d. On the effect of temporal scale on the model simulations

To evaluate the impact of temporal scale, we compare the mean and standard deviation of the generated fast and slow flow components before and after aggregation of the model output over monthly periods. Similar to the spatial aggregation, the summary metrics in Fig. 10 illustrate that the fast and slow reservoirs have more overlap for longer monthly aggregation periods.

e. VIC model fast and slow simulations across CAMELS dataset

We analyze the calibrated simulations of runoff and baseflow of the original VIC, VIC$_{original}$, in this study, for

the 531 catchments of CAMELS dataset for continental United States based on the work of Mizukami et al. (2017). We evaluate the mean and standard deviation of the fast and slow reacting reservoirs in every basin (Fig. 11a). Figure 11 clearly shows that the fast and slow reacting reservoirs have distinctive behavior, which is similar to the results from the single point simulations. We also evaluate the mean of the fast and slow reservoir in Fig. 11b, which shows that most of the model simulations for the basins are putting more emphasis on the baseflow generation than surface runoff.

Furthermore, Fig. 12a indicates the recession parameter a for observed streamflow has higher values compared with the same recession parameter for the VIC simulations as the cloud of points for the observation recession parameters are shifted to higher values comparably. The median values also shows a higher value for the recession parameter a for observed streamflow. Figure 12b also indicates that, overall, the recession
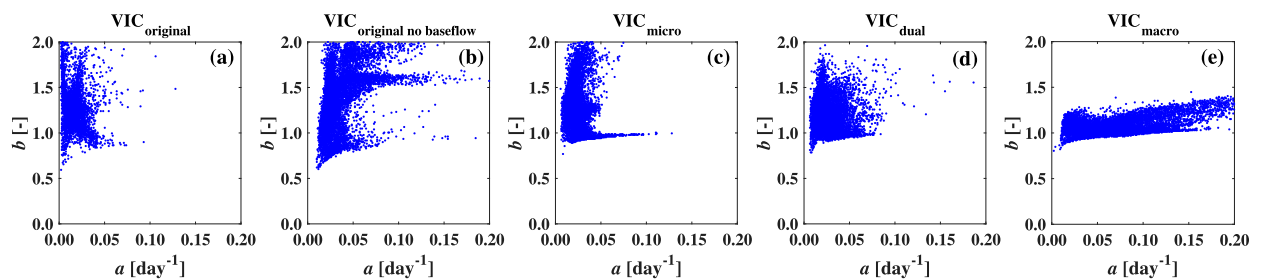

FIG. 8. The comparison of the recessions analysis, a and b parameters, for different VIC variations (a) VIC$_{original}$, (b) VIC$_{original\_no\_baseflow}$, (c) VIC$_{micro}$, (d)VIC$_{dual}$, and (e)VIC$_{macro}$.
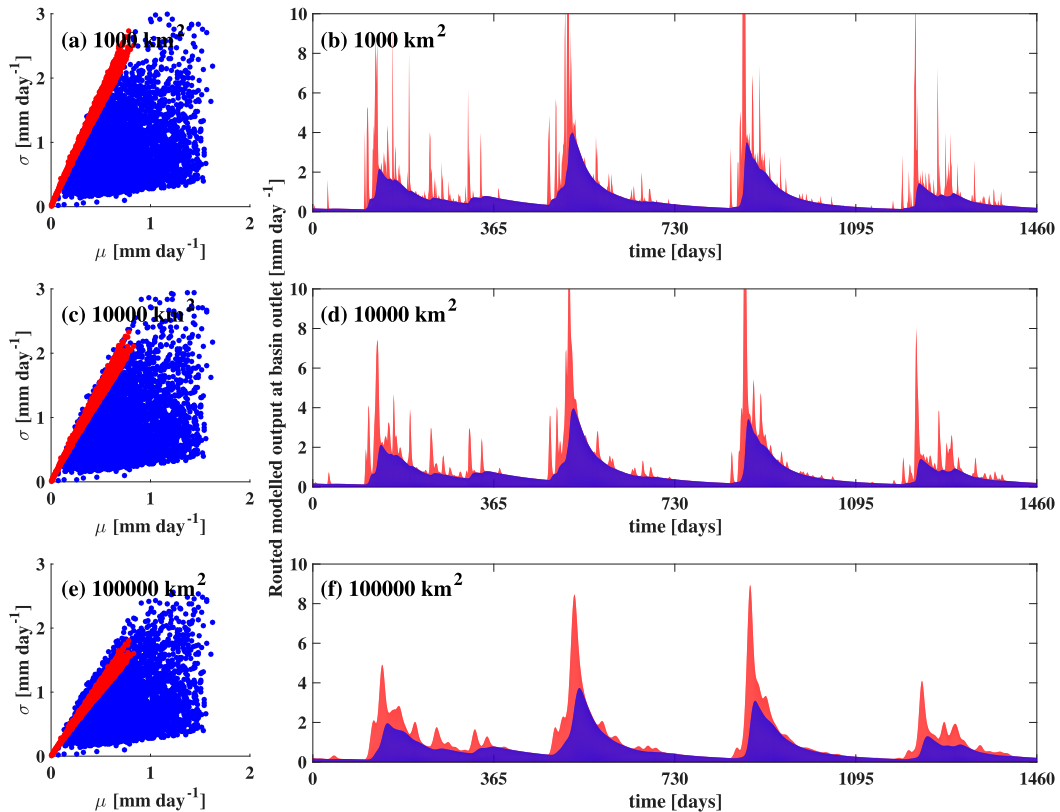
FIG. 9. Mean and standard deviation for basin sizes of (a) 1000, (c) 10 000, and (e) 100 000 km^2 and mean of the simulated fast and slow flow for basin sizes of (b) 1000, (d) 10 000, and (f) 100 000 km^2 for the simulation from original VIC, VIC$_{original}$.

parameter a is larger for the observation than the simulation, which reflects on the fact that model simulations are often simulating longer recessions than is usually observed in streamflow. It should be noted that the VIC simulations of the CAMELS dataset at the basin level underwent an in-basin routing that already masked the runoff and baseflow segregation of the VIC model, however the applied in-basin routing is not sufficiently enough to adjust the existing bias between the recession parameters in simulated and observed streamflow.

5. Discussion

In this study we showed that the structure of land models greatly simplifies the heterogeneity of the flow paths in nature. The land model signature is flashy runoff superimposed on slow recessions. Such behavior is constrained by the land model structure, imposing difficulties in simulating the temporal variability in streamflow across a range of different environments.

Such distinctive model behavior is masked by two major factors that can hamper the structural diagnosis of land model weaknesses. First, model simulations are

often aggregated over longer periods of time such as weekly or monthly (Melsen et al. 2016), which might hide the high temporal resolution deficiencies of simulations at the grid or point scale. Second, model simulations are typically aggregated over a large area using a routing scheme, which might also mask the subsurface structural deficiencies of the land models.
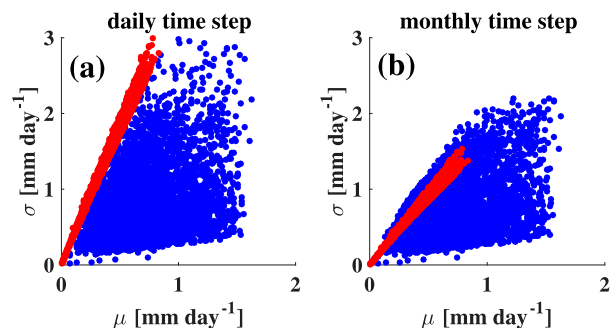

FIG. 10. The comparison between the mean and standard deviation of fast and slow components of VIC$_{original}$ in red and blue, respectively, aggregated at a (a) daily time step and (b) monthly time step.
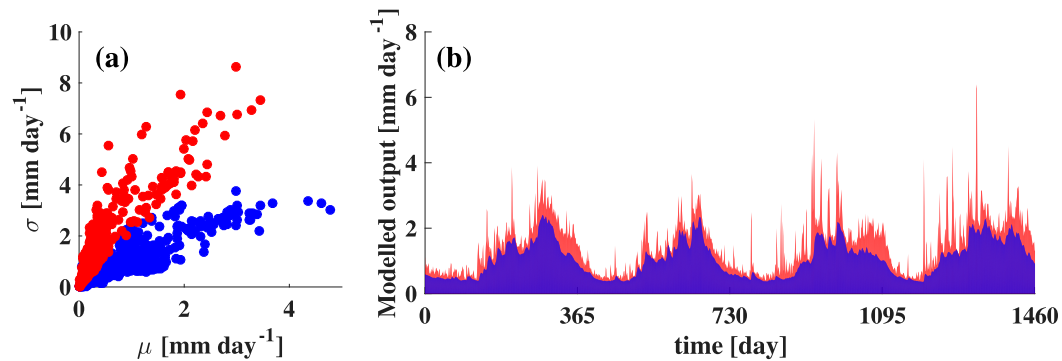
FIG. 11. (a) The mean and standard deviation (each point represents a CAMLES dataset basin) and (b) mean of the simulated runoff and baseflow for the original VIC setup, VIC$_{\text{original}}$, for all the 531 catchments of the CAMELS dataset over time across the continental United States.

This raises the question "Do land models get the right answers for the right reasons?" (Kirchner 2006). We argue that a structurally ill-posed model at the grid scale is able to yield satisfactory streamflow simulations if aggregated over spatiotemporal scales, yet ill-posed models are clearly poorly suited for many of their desired applications. The presented analyses on the adequacy of the land model structures can contribute to the ongoing debate on hyper resolution models (Wood et al. 2011; Beven and Cloke 2012; Wood et al. 2012; Beven et al. 2015). We argue that hyper resolution efforts need to fundamentally evaluate the ability of models in reproducing the temporal variability of streamflow across different space and time scales. Moreover, we emphasize the need for more diagnostic evaluation of land models in the context of flood forecasting. For example, and given our results, land models have difficulty simulating the rapidly rising limb of the hydrograph for shorter residence times. Additional model development is necessary for land models be applicable for a wider range of applications.

A possible way forward to include the preferential flow using the current land model structure, although the structure does not account for macropores explicitly, is to allow multiple landscape tiles with different soil properties (e.g., see Harman et al. 2009). This enables selecting soil parameters from a distribution rather than relying on a suggested value or only one value. Our case of the VIC$_{\text{macro}}$ model presented in section 3a(2), can be a specific case of the model with variable soil parameters: we can increase the fraction of preferential flow by assuming higher hydraulic conductivity than is suggested in laboratory soil properties and soil maps [for reading and references, refer to Nimmo (2012), Jarvis et al. (2016), and Beven (2018)]. Therefore, it is possible to mimic the VIC$_{\text{macro}}$ simulations using the original VIC model, VIC$_{\text{original}}$, with

two or more landscape tiles. Dual or multiple representations of the soil property provides the models with more agility to mimic the system behavior rather than choosing a fixed parameter. This is aligned with the concept of dual/multipermeability discussed widely in the context of water movement in the prose medium in modeling framework such as HYDRUS (Šimůnek and van Genuchten 2008). The dual/multipermeability is an attempt to achieve a more realistic representation of natural system behavior using the widely used Richards formulation for water movement in a porous medium to fulfill the observations that often do not follow the uniformity assumption. We remind the reader that identifying the dual or multiple characteristics of a porous medium may not be a straightforward task, as these data should be inferred from the observation and may depend on various factors especially at the larger scale, as it is often the case for land models. Therefore, identifying and inferencing the parameters and
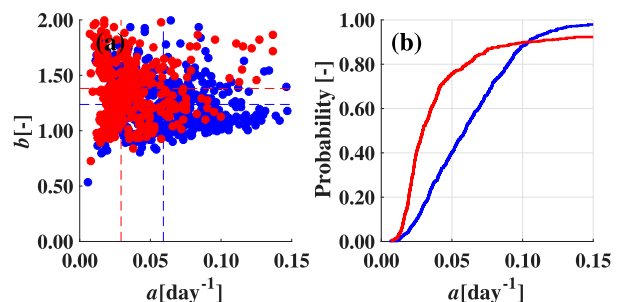

FIG. 12. The comparison of the recessions analysis, a and b parameters, for observed streamflow (blue) and VIC simulation (red) for 531 catchments of the CAMELS dataset across the continental United States. The dashed line indicates the median of the recession parameters for the observation and simulations in blue and red, respectively. (b) The cumulative distribution plot for the recession parameter a for observed and simulated streamflow in blue and red, respectively.

structure of macropore flow, due to heterogeneity and structure of the pore connectivity, is case and scale dependent. The future challenge remains on how to adequately understand and conceptualize the wide range of preferential flow types at various scales [e.g., rapid subsurface flow at hillslope scale or preferential recharge to (perched) aquifer] and how to identify the parameters of the designed models to achieve the highest fidelity. There have been studies that investigated the effect of explicitly representing preferential flow for hillslopes in land models (Soulis et al. 2000; Craig et al. 2010; Hazenberg et al. 2015), however including the preferential flow explicitly in the land models and its effect of the behavior of land models, preferential flow and soil moisture, is still in its infancy (e.g., see Milly et al. 2014; Nijzink et al. 2016; Rahman and Rosolem 2017).

6. Conclusions

The conclusions from this paper can be summarized as follows:

- The typical structure and formulation of land models does not allow for macropore water movement. The lack of macropore water movement results in two distinct fast and slow model simulations, characterized by flashy runoff superimposed on slow recessions.
- Without representation of macropore water movement in the model, the land model structure has difficulties simulating the range of natural system responses often observed in streamflow.
- The structural deficiencies and assumptions in land models are masked by aggregating results over longer time periods, such as month, or over larger areas through routing.

Based on the results presented in this paper, we join the call from Fan et al. (2019) to explicitly include more of the known processes at hillslope and catchment scales in land models. This is important because hillslopes are known to be fundamental elements of catchment response, especially during flood events. We therefore, encourage more diagnostic approaches to evaluate the utility of land models for flood forecasting. Representing preferential flow can open the door to more comprehensive model analyses (e.g., recession analyses) that can help identify model parameter values. An alternative path forward to include the macropores implicitly in the current land models is to allow a soil property distribution at a grid rather than fixing the soil parameters to perhaps highly uncertain soil maps and their attributes. Moreover, our proposed structural changes can better accommodate the two water world hypothesis

(McDonnell 2014), where explicitly representing preferential flow enables passing new water to the stream while older water remains in soil micropores. We look forward to more hydrologically oriented land model development, along with more incisive diagnostic studies to identify and address weaknesses in land models.

APPENDIX A

Toy Model

The water balance functions for the toy model are summarized as follows (see Fig. A1).

a. Snow

The snow reservoir S_{snow} is fed by q_{snow} (mm day^{-1}), which represents the precipitation as snow when air temperature T_{air} (°C) is less than T_{snow} (°C). Model parameter T_{snow} defines the temperature below which the precipitation is snow. The snow reservoir is depleted by snowmelt q_{melt} (mm day^{-1}), which is based on the degree-day factor c_{ddf} (mm °C^{-1} day^{-1}) that correlated the snowmelt q_{melt} to the temperature difference above T_{snow} ($T_{\text{air}} - T_{\text{snow}}$).

b. Unsaturated zone

The melted snow q_{melt} from the snow reservoir or rain q_{rain} partly enters the soil moisture reservoir. This partitioning is based on the amount of the water stored in the unsaturated zone S_u (mm) and two model parameters, the maximum storage of the unsaturated zone $S_{u\text{max}}$ (mm) and the shape function β (unitless). It is assumed that the snowmelt q_{melt} and rainfall P_{rain} entirely enters the unsaturated zone if the storage of the unsaturated zone is less than the fraction identified by f_{thr} that represents the system threshold for responding to precipitation based on the fill and spill mechanism. The unsaturated zone is depleted by two processes. Transpiration E_T (mm day^{-1}), which is calculated as a fraction of potential evaporation E_{pot} (mm day^{-1}), is calculated based on daily temperature using the Hamon equations (Hamon 1960). The transpiration is considered to be equal to potential evaporation when the amount of water stored in the soil moisture reservoir is more than a fraction of the total storage $S_{u\text{max}}$ described by f_{crit} (unitless).
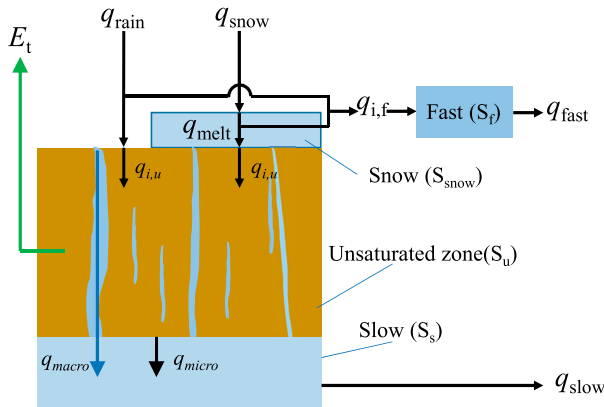
FIG. A1. The schematization for the toy models. By setting the macropore volume fraction D to 0, the slow reservoir S_s is only fed by micropore water movement q_{micro}, which is named as Model$_{micro}$. By setting the macropore volume fraction to a value between 0 and 1, the slow reservoir is fed by both macropore and micropore water movement, q_{micro} and q_{macro}, which is named as Model$_{dual}$. By setting the k_{micro} to 0 and allowing the macropore volume fraction between 0 and 1, the slow reservoir is fed only by macropore water movement q_{macro}, which is named as Model$_{macro}$.

Transpiration is reduced linearly with the soil moisture when the soil moisture falls below the fraction identified by f_{crit}. The soil moisture is also depleted by the percolation to slow or saturated reservoir. The percolation is linearly related to the level of saturation, defined as the ratio of soil moisture to the maximum soil moisture $S_u/S_{u\max}$ and the maximum percolation rate or saturated hydraulic conductivity defined by k_{micro} (mm day^{-1}).

c. Fast and slow saturated zones

The amount of water that did not enter into the soil moisture will be split between the fast and slow reservoir, S_F and S_S, based on macropore volume fraction D (unitless) that represents the fraction of macropore flow to the slow saturated zone. The fraction of the water that does not enter the soil moisture in the slow reservoir is considered to be surface runoff and is added to the fast reservoir. The fast and slow reservoirs are linear reservoirs that are defined by parameters k_{fast} (day^{-1}) and k_{slow} (day^{-1}), respectively. The simulated discharge is the summation of the fast and slow flows, $Q_m = q_{slow} + q_{fast}$.

The model state equations in Table A1 are solved using a forward explicit method in which the states of the next time step are dependent only on the state of the current time step of the model.

APPENDIX B

The Synthetic Routing Scheme to Evaluate the Effect of Routing on Hydrograph Simulation

To evaluate the effect of the routing scheme on the simulated hydrograph at the outlet of a basin, a synthetic routing scheme is designed as below:

1) We assume a basin total area (e.g., 1000 km^2).
2) The total number of grid cells are the total area of basin divided by 100 km^2 (assuming that each grid is 10 km $\times$ 10 km).

TABLE A1. The water balance and flux equations for the toy model.

Reservoir	Water balance equation	Fluxes
Snow	$dS_{snow}/dt = q_{snow} - q_{melt}$	$q_{snow} = \begin{cases} 0, & T_{air} > T_{snow} \\ P, & T_{air} \le T_{snow} \end{cases}$
		$q_{melt} = \begin{cases} c_{ddf}(T_{air} - T_{snow}), & T_{air} > T_{snow} \\ 0, & T_{air} \le T_{snow} \end{cases}$
Unsaturated	$dS_u/dt = q_{i,u} - q_{micro} - E_T$	$q_{i,u} = q_{total}(1 - A_s)$
		$E_T = \begin{cases} E_{pot}, & S_u > f_{crit}S_{u\max} \\ E_{pot}[S_u/(f_{crit}S_{u\max})], & S_u \le f_{crit}S_{u\max} \end{cases}$
		$q_{micro} = k_{micro}(S_u/S_{u\max})$
		$A_s = \begin{cases} 0, & S_u \le f_{thr}S_{u\max} \\ [(S_u - f_{thr}S_{u\max})/(S_{u\max} - f_{thr}S_{u\max})]^\beta, & S_u > f_{thr}S_{u\max} \end{cases}$
		$q_{rain} = \begin{cases} P, & T_{air} > T_{snow} \\ 0, & T_{air} \le T_{snow} \end{cases}$
		$q_{total} = q_{melt} + q_{rain}$
Fast	$dS_f/dt = q_{i,f} - q_{fast}$	$q_{i,f} = q_{total}A_s(1 - D)$
		$q_{fast} = k_{fast}S_f$
Slow	$dS_s/dt = q_{i,s} - q_{slow}$	$q_{i,s} = q_{macro} + q_{micro}$
		$q_{slow} = k_{slow}S_s$
		$q_{macro} = q_{total}A_sD$

3) The longest river in the basin can be correlated with the basin size using Hack's law (Hack 1957):

$$L_{\max} = 1.4A^{0.6},$$

in which A is the basin area in square miles and $L_{\max}$ is the longest river segment in the basin expressed in miles.

4) We assume that the grid cells have a distance to the outlet that follows a normal distribution with a mean of half of the maximum length obtained from the Hack's law and a standard deviation equals 1/3 of the mean value.

5) We assume that the convolution function for every grid cell follows the following formulas:

$$h(x,t) = \frac{x}{2t\sqrt{\pi t D_{\mathrm{diff}}}} \exp\left[-\frac{(Vt-x)^2}{4D_{\mathrm{diff}}t}\right],$$

in which D_{diff} is diffusivity that is fixed at $4000\,\mathrm{m}^2\,\mathrm{s}^{-1}$ and V is velocity of $0.5\,\mathrm{m\,s}^{-1}$. Parameter $h(x,t)$ represents the routed discharge at point of x and time t from every single grid cell.

6) The final simulated hydrograph is sum of the convoluted grid simulation at the outlet.

REFERENCES

Addor, N., A. J. Newman, N. Mizukami, and M. P. Clark, 2017: The CAMELS data set: Catchment attributes and meteorology for large-sample studies. *Hydrol. Earth Syst. Sci.*, **21**, 5293–5313, https://doi.org/10.5194/hess-21-5293-2017.

Archfield, S. A., and Coauthors, 2015: Accelerating advances in continental domain hydrologic modeling. *Water Resour. Res.*, **51**, 10 078–10 091, https://doi.org/10.1002/2015WR017498.

Beven, K., 2018: A century of denial: preferential and non-equilibrium water flow in soils, 1864-1984. *Vadose Zone J.*, **17**, 180153, https://doi.org/10.2136/vzj2018.08.0153.

Beven, K. J., and H. L. Cloke, 2012: Comment on "Hyper-resolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water" by Eric F. Wood et al. *Water Resour. Res.*, **48**, W01801, https://doi.org/10.1029/2011WR010982.

Beven, K., and P. Germann, 1982: Macropores and water flow in soils. *Water Resour. Res.*, **18**, 1311–1325, https://doi.org/10.1029/WR018i005p01311.

——, H. Cloke, F. Pappenberger, R. Lamb, and N. Hunter, 2015: Hyperresolution information and hyperresolution ignorance in modelling the hydrology of the land surface. *Sci. China Earth Sci.*, **58**, 25–35, https://doi.org/10.1007/s11430-014-5003-4.

Boone, A., and P. J. Wetzel, 1996: Issues related to low resolution modeling of soil moisture: Experience with the PLACE model. *Global Planet. Change*, **13**, 161–181, https://doi.org/10.1016/0921-8181(95)00044-5.

Clark, M. P., D. E. Rupp, R. A. Woods, H. J. Tromp-van Meerveld, N. E. Peters, and J. E. Freer, 2009: Consistency between hydrological models and field observations: Linking processes at the hillslope scale to hydrological responses at the watershed scale. *Hydrol. Processes*, **23**, 311–319, https://doi.org/10.1002/hyp.7154.

——, and Coauthors, 2015: Improving the representation of hydrologic processes in Earth system models. *Water Resour. Res.*, **51**, 5929–5956, https://doi.org/10.1002/2015WR017096.

Craig, J. R., G. Liu, and E. D. Soulis, 2010: Runoff–infiltration partitioning using an upscaled Green–Ampt solution. *Hydrol. Processes*, **24**, 2328–2334, https://doi.org/10.1002/hyp.7601.

Cuntz, M., J. Mai, L. Samaniego, M. Clark, V. Wulfmeyer, O. Branch, S. Attinger, and S. Thober, 2016: The impact of standard and hard-coded parameters on the hydrologic fluxes in the Noah-MP land surface model. *J. Geophys. Res. Atmos.*, **121**, 10 676–10 700, https://doi.org/10.1002/2016JD025097.

Davies, J., K. Beven, A. Rodhe, L. Nyberg, and K. Bishop, 2013: Integrated modeling of flow and residence times at the catchment scale with multiple interacting pathways. *Water Resour. Res.*, **49**, 4738–4750, https://doi.org/10.1002/wrcr.20377.

Fan, Y., and Coauthors, 2019: Hillslope hydrology in global change research and Earth system modeling. *Water Resour. Res.*, **55**, 1737–1772, https://doi.org/10.1029/2018WR023903.

Hack, J. T., 1957: Studies of longitudinal stream profiles in Virginia and Maryland. Geological Survey Professional Paper 294-B, 97 pp., https://pubs.usgs.gov/pp/0294b/report.pdf.

Hamman, J. J., B. Nijssen, T. J. Bohn, D. R. Gergel, and Y. Mao, 2018: The Variable Infiltration Capacity model version 5 (VIC-5): Infrastructure improvements for new applications and reproducibility. *Geosci. Model Dev.*, **11**, 3481–3496, https://doi.org/10.5194/gmd-11-3481-2018.

Hamon, W. R., 1960: Estimating potential evapotranspiration. B.S. dissertation, Dept. of Civil and Sanitary Engineering, Massachusetts Institute of Technology, 75 pp., https://dspace.mit.edu/handle/1721.1/79479.

Harman, C. J., M. Sivapalan, and P. Kumar, 2009: Power law catchment-scale recessions arising from heterogeneous linear small-scale dynamics. *Water Resour. Res.*, **45**, W09404, https://doi.org/10.1029/2008WR007392.

Hazenberg, P., Y. Fang, P. Broxton, D. Gochis, G. Y. Niu, J. D. Pelletier, P. A. Troch, and X. Zeng, 2015: A hybrid-3D hillslope hydrological model for use in Earth system models. *Water Resour. Res.*, **51**, 8218–8239, https://doi.org/10.1002/2014WR016842.

Jarvis, N., J. Koestel, and M. Larsbo, 2016: Understanding preferential flow in the vadose zone: Recent advances and future prospects. *Vadose Zone J.*, **15** (12), https://doi.org/10.2136/vzj2016.09.0075.

Kirchner, J. W., 2006: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resour. Res.*, **42**, W03S04, https://doi.org/10.1029/2005WR004362.

Koide, S., and H. S. Wheater, 1992: Subsurface flow simulation of a small plot at Loch Chon, Scotland. *Hydrol. Processes*, **6**, 299–326, https://doi.org/10.1002/hyp.3360060306.

Koster, R. D., and S. P. P. Mahanama, 2012: Land surface controls on hydroclimatic means and variability. *J. Hydrometeor.*, **13**, 1604–1620, https://doi.org/10.1175/JHM-D-12-050.1.

——, and P. C. D. Milly, 1997: The interplay between transpiration and runoff formulations in land surface schemes used with atmospheric models. *J. Climate*, **10**, 1578–1591, https://doi.org/10.1175/1520-0442(1997)010<1578:TIBTAR>2.0.CO;2.

Lee, D. H., and L. M. Abriola, 1999: Use of the Richards equation in land surface parameterizations. *J. Geophys. Res.*, **104**, 27 519–27 526, https://doi.org/10.1029/1999JD900951.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges, 1994: A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res.*, **99**, 14 415–14 428, https://doi.org/10.1029/94JD00483.

Manabe, S., 1969: Climate and the ocean circulation: I. The atmospheric circulation and the hydrology of the earth's surface. *Mon. Wea. Rev.*, **97**, 739–774, https://doi.org/10.1175/1520-0493(1969)097<0739:CATOC>2.3.CO;2.

McDonnell, J. J., 2014: The two water worlds hypothesis: Ecohydrological separation of water between streams and trees? *Wiley Interdiscip. Rev.: Water*, **1**, 323–329, https://doi.org/10.1002/wat2.1027.

Melsen, L. A., A. J. Teuling, P. J. Torfs, R. Uijlenhoet, N. Mizukami, and M. P. Clark, 2016: HESS Opinions: The need for process-based evaluation of large-domain hyperresolution models. *Hydrol. Earth Syst. Sci.*, **20**, 1069–1079, https://doi.org/10.5194/hess-20-1069-2016.

Mendoza, P. A., M. P. Clark, M. Barlage, B. Rajagopalan, L. Samaniego, G. Abramowitz, and H. Gupta, 2015: Are we unnecessarily constraining the agility of complex process-based models? *Water Resour. Res.*, **51**, 716–728, https://doi.org/10.1002/2014WR015820.

Milly, P. C., and Coauthors, 2014: An enhanced model of land water and energy for global hydrologic and earth-system studies. *J. Hydrometeor.*, **15**, 1739–1761, https://doi.org/10.1175/JHM-D-13-0162.1.

Mirus, B. B., and J. R. Nimmo, 2013: Balancing practicality and hydrologic realism: A parsimonious approach for simulating rapid groundwater recharge via unsaturated-zone preferential flow. *Water Resour. Res.*, **49**, 1458–1465, https://doi.org/10.1002/wrcr.20141.

——, B. A. Ebel, C. S. Heppner, and K. Loague, 2011: Assessing the detail needed to capture rainfall-runoff dynamics with physics-based hydrologic response simulation. *Water Resour. Res.*, **47**, W00H10, https://doi.org/10.1029/2010WR009906.

Mizukami, N., M. P. Clark, A. J. Newman, A. W. Wood, E. D. Gutmann, B. Nijssen, O. Rakovec, and L. Samaniego, 2017: Towards seamless large-domain parameter estimation for hydrologic models. *Water Resour. Res.*, **53**, 8020–8040, https://doi.org/10.1002/2017WR020401.

Nijzink, R. C., and Coauthors, 2016: The importance of topography-controlled sub-grid process heterogeneity and semi-quantitative prior constraints in distributed hydrological models. *Hydrol. Earth Syst. Sci.*, **20**, 1151–1176, https://doi.org/10.5194/hess-20-1151-2016.

Nimmo, J. R., 2012: Preferential flow occurs in unsaturated conditions. *Hydrol. Processes*, **26**, 786–789, https://doi.org/10.1002/hyp.8380.

——, and L. Mitchell, 2013: Predicting vertically nonsequential wetting patterns with a source-responsive model. *Vadose Zone J.*, **12** (4), https://doi.org/10.2136/vzj2013.03.0054.

Rahman, M., and R. Rosolem, 2017: Towards a simple representation of chalk hydrology in land surface modelling. *Hydrol. Earth Syst. Sci.*, **21**, 459–471, https://doi.org/10.5194/hess-21-459-2017.

Šimůnek, J., and M. T. van Genuchten, 2008: Modeling nonequilibrium flow and transport processes using HYDRUS. *Vadose Zone J.*, **7**, 782–797, https://doi.org/10.2136/vzj2007.0074.

Soulis, E. D., K. R. Snelgrove, N. Kouwen, F. Seglenieks, and D. L. Verseghy, 2000: Towards closing the vertical water balance in Canadian atmospheric models: Coupling of the land surface scheme CLASS with the distributed hydrological model WATFLOOD. *Atmos.–Ocean*, **38**, 251–269, https://doi.org/10.1080/07055900.2000.9649648.

USGS, 2019: USGS Water Data for the Nation. National Water Information System, accessed 30 November 2019, https://doi.org/10.5066/F7P55KJN.

Vogel, R. M., and C. N. Kroll, 1992: Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics. *Water Resour. Res.*, **28**, 2451–2458, https://doi.org/10.1029/92WR01007.

Weedon, G. P., G. Balsamo, N. Bellouin, S. Gomes, M. J. Best, and P. Viterbo, 2014: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resour. Res.*, **50**, 7505–7514, https://doi.org/10.1002/2014WR015638.

Wheater, H. S., S. Tuck, R. C. Ferrier, A. Jenkins, F. M. Kleissen, T. A. B. Walker, and M. B. Beck, 1993: Hydrological flow paths at the Allt a'Mharcaidh catchment: An analysis of plot and catchment scale observations. *Hydrol. Processes*, **7**, 359–371, https://doi.org/10.1002/hyp.3360070403.

Wood, E. F., D. P. Lettenmaier, and V. G. Zartarian, 1992: A land-surface hydrology parameterization with subgrid variability for general circulation models. *J. Geophys. Res.*, **97**, 2717–2728, https://doi.org/10.1029/91JD01786.

——, and Coauthors, 2011: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. *Water Resour. Res.*, **47**, W05301, https://doi.org/10.1029/2010WR010090.

——, and Coauthors, 2012: Reply to comment by Keith J. Beven and Hannah L. Cloke on "Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water". *Water Resour. Res.*, **48**, W01802, https://doi.org/10.1029/2011WR011202.

Zhao, R. J., Y. L. Zhang, L. R. Fang, X. R. Liu, and Q. S. Zhang, 1980: The Xinanjiang model. *IAHS Publ.*, **129**, 351–356.