



LM4-SHARC v1.0: resolving the catchment-scale soil–hillslope aquifer–river continuum for the GFDL Earth system modeling framework

Minki Hong¹, Nathaniel Chaney², Sergey Malyshev³, Enrico Zorzetto⁴, Anthony Preucil³, and Elena Shevliakova³

¹Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA

²Department of Civil and Environmental Engineering, Duke University, Durham, NC, USA

³NOAA OAR Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA

⁴Department of Earth and Environmental Science, New Mexico Institute of Mining and Technology, Socorro, NM, USA

Correspondence: Minki Hong (mh0093@princeton.edu)

Received: 30 June 2024 – Discussion started: 18 September 2024

Revised: 14 January 2025 – Accepted: 27 January 2025 – Published: 14 April 2025

Abstract. Catchment-scale representation of the groundwater and its interaction with other parts of the hydrologic cycle is crucial for accurately depicting the land water–energy balance in Earth system models (ESMs). Despite existing efforts to describe the groundwater in the land component of ESMs, most ESMs still need a prognostic framework for describing catchment-scale groundwater based on its emergent properties to understand the implications for the broader Earth system. To fill this gap, we developed a new parameterization scheme to resolve the groundwater and its two-way interactions with the unsaturated soil and stream at the catchment scale. We implemented this new parameterization scheme (SHARC, or the soil–hillslope aquifer–river continuum) in the Geophysical Fluid Dynamics Laboratory (GFDL) land model (i.e., LM4-SHARC) and evaluated its performance. By bridging the gap between hydraulic groundwater theory and ESM land hydrology, the new LM4-SHARC provides a path to learning groundwater emergent properties from available streamflow data (i.e., recession analysis), enhancing the representation of subgrid variability in water–energy states induced by the groundwater. LM4-SHARC has been applied to the Providence headwater catchment at Southern Sierra, NV, and tested against in situ observations. We found that LM4-SHARC leads to noticeable improvements in the representation of key hydrologic variables such as streamflow, near-surface soil moisture, and soil temperature. In addition to enhancing the representation of the water and energy balance, our analysis showed that accounting for groundwater

convergence can induce a more significant hydrologic contrast, with higher sensitivity of soil water storage to groundwater properties in the riparian zone. Our findings indicate the feasibility of incorporating two-way interactions among groundwater, unsaturated soil, and streams into the hydrological components of ESMs and show a further need to explore the implications of these interactions in the context of Earth system dynamics.

1 Introduction

The significance of understanding the relationship between the hydrologic cycle and climate variability has been increasingly recognized in a warming climate (Milly et al., 2008). As a major component of the terrestrial water cycle, groundwater (i.e., water in saturated zones beneath the land surface) plays a pivotal role in developing the surface thermal energy and moisture dynamics and the land–atmosphere coupling by regulating how the water and thermal energy are stored and transported across the landscape (Andreae et al., 2002; Mu et al., 2011; McCabe et al., 2008; Gentine et al., 2019; Fan, 2015). The effects of the groundwater state (e.g., storage) on the water–energy balance at the land surface have been discussed in several studies based on an explicitly treated groundwater scheme (Miguez-Macho et al., 2007; Liang et al., 2003; Yeh and Eltahir, 2005; Zeng et al., 2016; Maxwell et al., 2011; Maxwell and Kollet, 2008). The stud-

ies commonly found higher sensitivity of the surface energy balance to groundwater storage if the water table is shallow (e.g., the water table depth is less than 5 m).

However, many Earth system models (ESMs) only represent the top few meters of soil, and most, if not all, ignore two-way interactions between the groundwater and other components of the hydrologic cycle. In particular, ESMs lack proper land surface scheme(s) to represent the two-way interactions between the groundwater and the stream/river while conserving the hydraulic continuity between them. In most cases, the river routing module implemented in the ESMs exists mainly for traditional purposes (e.g., linking precipitation-induced runoff to the ocean) without the ability to consider the groundwater variations driven by the horizontal hydraulic gradients between the stream and the groundwater (Li et al., 2013; De Rosnay et al., 2002; Lawrence et al., 2011, 2019; Miguez-Macho and Fan, 2012; Best et al., 2011; Takata et al., 2003).

Furthermore, the resolution of climate-/water-related information provided by the current ESMs is too low to characterize hydrological extremes and address the stakeholders' need to evaluate potential impact of future climate change. Even atmosphere–land water information produced at a resolution of 0.25° (Delworth et al., 2020) – the highest resolution in ESMs – is still not sufficiently fine to generate fine-scale data for decision-making. The land components of ESMs generate the streamflow estimates only at a grid scale (typically ranging from 0.5 to 1.0°) (Oleson et al., 2013; Miguez-Macho et al., 2007; Fan et al., 2007; Pappenberger et al., 2012; Campoy et al., 2013), and since such coarse-resolution streamflow data, for example, is not suitable for directly locating hydrologic events at a fine scale, even in the case of extremes (e.g., flooding), the scale mismatch (between the stakeholders and ESMs) could reduce the usability of the ESM-derived projections. Another challenge lies in the high degree of land heterogeneity with respect to soil and topographic characteristics, which needs to be resolved at the ESM grid scale. Considering the significant impacts of the fine-scale soil/topographic properties on the hydrologic processes, the two-way interactions between the groundwater and the rest of the hydrologic cycle must be parameterized at the subgrid scale (e.g., catchment) to adequately represent the subgrid variability in hydrologic states and to show its implications for the broader Earth system processes and interactions (Gleeson et al., 2020; Xu et al., 2023; Maxwell and Kollet, 2008; Pokhrel et al., 2013).

Catchments provide an appropriate scale to properly capture such subgrid spatial heterogeneity and its effect on the interactions among the hydrological cycle components (Clark et al., 2015; Fan et al., 2019; Blyth et al., 2021). This is mainly due to the suitability of the catchment scale in describing the topographically driven flow characteristics of water and other major fluxes (e.g., thermal energy) caused by water transport. In fact, catchments are sometimes considered hydrologic spatial units (or response units), where the

theoretical conceptualization of surface and subsurface water transport can be tested with readily available observational data, such as streamflow measurements (Sivapalan, 2006; Troch et al., 2013; Kirchner, 2009). Over the past decades, a strong effort has been made to depict the hydraulically based interactions among the distinct flow domains at the hillslope and catchment scales (Kollet and Maxwell, 2006; Niu et al., 2011; Shen and Phanikumar, 2010; Gochis, 2021). This effort mainly aimed to represent the water–energy coupled balance while considering the time-dependent nonlinear relationship between water states and fluxes based on Darcian flow (Clark et al., 2015; Fan et al., 2019).

However, when specifically targeting the representation of the groundwater and its interactions with the overlying soil and the stream/river, the spatial heterogeneity of groundwater properties (e.g., diffusivity or effective (drainable) porosity) remains a significant challenge for enhancing the predictability of terrestrial water storage and exchange fluxes (Clark et al., 2009; Jing et al., 2019). Notably, accurate description of groundwater properties/processes at the catchment-scale is crucial because the hydrologic convergence (to the riparian zone) and divergence (from the hilltops) by the groundwater movement significantly affects the catchments' water and energy dynamics (Fan, 2015; Miguez-Macho and Fan, 2012; Chen and Hu, 2004; Maxwell et al., 2007). Even in previous studies that captured the interactions among the soil, groundwater, and streams, the properties of the groundwater were treated as constant and not as emergent and were treated as dynamic properties as the result of transient groundwater storage affected by the climate (e.g., precipitation) or human activities (e.g., land use change or groundwater pumping) (Li and Ameli, 2022; Bart and Hope, 2014; Jachens et al., 2020b; Hong and Mohanty, 2023a; Trotter et al., 2024). For example, these studies empirically fit coefficients to a storage–discharge function (e.g., bucket model) or to a predefined soil properties dataset to parameterize the groundwater domains (Zeng et al., 2016; Gochis, 2021; Newman et al., 2014; Kollet and Maxwell, 2008; Leung et al., 2011; Li and Ameli, 2022). Thus, a theoretical approach to capture the dynamic, emergent properties of the catchment-scale groundwater has largely been absent in these efforts.

The Dupuit–Forchheimer (DF) approximation allows the incorporation of the catchment-scale groundwater into ESMs and accounts for emergent properties/processes that shape the interactions with the unsaturated soil and the stream (Dupuit, 1863; Forchheimer, 1986). According to the DF approximation, heterogeneity in groundwater properties can be represented by effective parameters reflecting the combined effects of the groundwater property variability on fluxes (e.g., baseflow), assuming that the lateral groundwater discharge is from the homogeneous/isotropic (or systemically defined) groundwater (Rupp and Selker, 2006; Strack, 1995; Beven and Kirkby, 1979). The DF approximation is especially relevant, as the heterogeneity of groundwater properties that are not representable below a model-specific resolution can be

considered one of the challenges for the accuracy of modeled data (Baroni et al., 2019; Maxwell et al., 2015; Niu et al., 2011; Bisht et al., 2017). The DF approximation is derived by recasting the analytical solution of the Boussinesq equation (Boussinesq, 1903) into the power law streamflow recession model by Brutsaert and Nieber (1977). This method of streamflow recession analysis provides a theoretical basis for estimating the effective groundwater properties on a catchment scale using (relatively readily) available streamflow observations (Tashie et al., 2021; Hong and Mohanty, 2023a; Vannier et al., 2014; Brutsaert and Lopez, 1998; Troch et al., 1993; Xu et al., 2018).

The catchment-scale processes can be resolved at the ESM grid scale via approaches to the subgrid model structure such as HydroBlocks (Chaney et al., 2016, 2018). The hierarchical multivariate clustering (HMC) approach, as implemented in LM4-HydroBlocks (Chaney et al., 2018), enables the partitioning of a macroscale land domain grid (e.g., 0.5° , 1.0°) into subgrid terrain units (termed tiles) (Milly et al., 2014; Zhao et al., 2018; Dunne et al., 2020). This study presents a new parameterization of the two-way soil–hillslope aquifer–river continuum, namely SHARC (v1.0), tailored to the fourth generation of the Geophysical Fluid Dynamics Laboratory (GFDL) land model (i.e., LM4), aiming to represent the catchment-scale two-way interactions among the unsaturated soil (i.e., vadose zone), the groundwater, and the stream/river in the GFDL Earth system modeling framework. LM4-SHARC has been developed by extending the hillslope hydrology scheme that is currently used in LM4-HydroBlocks (Subin et al., 2014; Chaney et al., 2018) and by relying on the catchment-scale hydrologic partitioning (of the macroscale grid cell) using HMC approach for the explicit treatment of divergence fluxes among soil, groundwater, and streams.

In sum, LM4-SHARC (1) explicitly characterizes the catchment-scale groundwater while accounting for its emergent properties, such as groundwater diffusivity; (2) represents the two-way water and energy exchanges in the vertical direction between the soil and the groundwater and in the lateral direction between the groundwater and the stream; and (3) accounts for the groundwater-induced variations in surface water–energy budgets to enhance the ESM’s land hydrology realism. This study primarily compares the two model configurations to evaluate the importance of accurately describing the catchment-scale groundwater in the land component of the ESM. As a proof of concept, we apply LM4-SHARC and LM4-HydroBlocks to a headwater catchment in the Sierra National Forest, Nevada, and evaluate the modeled outputs vs. the corresponding observations. Additionally, we discuss how the hydrologic convergence of liquid water and thermal energy due to the groundwater flow to the river valley (and divergence from the hilltop) could contribute to hydrologic contrast in a catchment based on simulations.

2 Methods

2.1 Characterization of the catchment-scale spatial domain and its heterogeneity

A hierarchical multivariate clustering (HMC) method (Chaney et al., 2018, 2016) was used to characterize the catchment-scale model domain and its spatial heterogeneity. The conceptual approach uses available environmental datasets (such as soil properties, topography, meteorology, and land cover) to characterize the subgrid spatial heterogeneity within each grid cell of an Earth system model (ESM) land domain (Chaney et al., 2018). In fact, the HMC method has been used previously when constructing land domains for LM4-HydroBlocks to account for the subgrid spatial variability in land properties within a regular latitude–longitude climate grid cell (e.g., at a 0.5° or 1.0° resolution). The land fraction of each grid cell in the LM4-HydroBlocks domain is partitioned into soil, glacier, and lake components. The soil component of a grid cell is composed of hillslopes clustered into a set of k characteristic hillslopes (CHs). Each CH has unique attributes, such as slope, aspect, and convexity, which are local averages of these fields obtained by high-resolution datasets. Each CH is further partitioned into l units denoted height bands (HBs), which are obtained by partitioning each CH based on elevation bins of size dh . Finally, each HB can be further divided into p clusters (i.e., tiles), and then each tile is considered a representative element volume (REV) of the soil with respect to the covariates incorporated for clustering (Chaney et al., 2018). The channels were delineated using an area threshold of $100\,000\text{ m}^2$. Three parameters, k , dh , and p , are required in the land pre-processing code: k determines the number of CHs in the land model domain, and dh defines the height difference between adjacent HBs, thus determining the number of HBs in a CH. The parameter p sets the number of tiles in an HB (i.e., the intra-HB variability). For further details on the HMC algorithm, see Chaney et al. (2018, 2021).

Unlike the previous configuration of the land input dataset for a regular grid cell, this study, for the first time, refined the existing HMC framework to generate a land input dataset for a catchment domain with an irregular boundary. This refinement has been designed to assess the GFDL land model’s performance at the catchment level while enhancing the interpretability of the land model’s hydrologic outputs. The catchment’s boundary was determined by computing the geographic extent of the area that drains to a specific point (i.e., the catchment outlet). Since the catchment areas are significantly variable (ranging from less than $\sim 1\text{ km}^2$ to hundreds of km^2), the resolution of the digital elevation model (DEM) used to cluster the terrain characteristics should be fine enough to capture the intra-catchment heterogeneity of terrain characteristics adequately. To this end, the USGS 3DEP 0.33 arcsec (i.e., 10 m resolution) DEM was used in this study (USGS, 2019).

2.2 Representation of the soil–groundwater processes in LM4-HydroBlocks

In the current configuration of GFDL LM4-HydroBlocks, the impermeable bedrock is assumed to exist 10 m below the ground, and the subgrid-scale (e.g., reach) river dynamics are not represented. The lateral liquid water flux between the saturated soil layers in each adjacent tile is considered groundwater flow. Hence, the groundwater properties depend on the surface/near-surface properties from which each soil column's soil properties are derived (Milly et al., 2014; Shevliakova et al., 2024). Each soil column in a catchment is composed of a vertical 1D model (i.e., soil–bedrock column) from the canopy air down to an impermeable bedrock layer (Milly et al., 2014). The processes resolved by the 1D model include the surface energy balance, vegetation dynamics, plant hydraulics, photosynthesis, snow physics, and soil thermal and hydraulic physics (Zorzetto et al., 2023; Subin et al., 2014; Milly et al., 2014). All the soil–bedrock columns are simulated with a 10 m soil depth at a 30 min physical time step. The 10 m soil–bedrock column is discretized into 20 layers. At the surface, the water flux boundary condition is tentatively set to the difference between the sum of rainfall and snowmelt minus evaporation at each time step (i.e., time-dependent flow). The heat flux boundary condition at the surface is determined by the balance between the turbulent and radiative fluxes. At the bottom of the soil–bedrock column (i.e., 10 m below the land surface), water flux is assumed to be zero, implying that the impermeable bedrock is always located 10 m below the land surface. Due to this impermeable layer assumption, the water table is always identified within 10 m of the land surface. A constant geothermal heat flux is prescribed at the bottom of the soil–bedrock columns (Milly et al., 2014).

The catchment-scale hydrologic structure within the macroscale grid cell is described based on the inter-tile connection enforced through the flow accumulation area derived from the digital elevation model (DEM) (Chaney et al., 2018). The water fluxes between adjacent tiles (i.e., inter-tile) are simulated on a layer-to-layer basis, defined following the soil layer order from the land surface according to the gradient of the total pressure head between the corresponding soil layers. The inter-tile heat and dissolved organic carbon fluxes advected by water transport are also represented. While the tile located right next to the river reach interacts with the river, the flux from the tile to the river is a one-way flux, as the reach-scale streams' (corresponding to a catchment) hydrographs are not accounted for in LM4-HydroBlocks. With respect to the flux exchange between the river and the soil–bedrock column (adjacent to the reach), the river stage is set to zero (static pressure head in the reach). Consequently, the water flux is invariably one-way, from the hillslope to the river.

2.3 Representation of the soil–hillslope aquifer–river interactions in LM4-SHARC

2.3.1 Two-way water transport and conservation in LM4-SHARC

LM4-SHARC solves the nonlinear Boussinesq equation, which was derived from the DF approximation to represent the lateral groundwater discharge flux from the hilltop to the river in homogeneous and isotropic unconfined groundwater (i.e., the stream–hillslope interface) (Basha, 2013; Hong and Mohanty, 2023b; Hornberger and Remson, 1970). According to the DF approximation, the unconfined groundwater flows laterally, and the lateral discharge flux is proportional to the saturated groundwater thickness (Dupuit, 1863; Forchheimer, 1986). Therefore, the lateral hydraulic gradient is the only driver of groundwater lateral discharge fluxes, as the equipotential lines in the saturated zone are assumed to be vertical (i.e., hydrostatic). For saturated groundwater flow in unconfined groundwater lying over an impermeable bedrock of slope θ , the lateral groundwater discharge flux is estimated following Eq. (1).

$$q_l = -K_s \left[\cos\theta \left(\frac{\partial N}{\partial x} \right) + \sin\theta \right], \quad (1)$$

where q_l is the speed of the lateral groundwater divergence flux (mm s^{-1}). K_s is the saturated lateral hydraulic conductivity (mm s^{-1}). Thus, the groundwater flow rate per unit width of the groundwater is given by $q_l N$, where N is the thickness of the groundwater layer perpendicular to the impermeable bedrock (m). Inserting flux Eq. (1) into the mass continuity equation yields the Boussinesq groundwater Eq. (2).

$$f \frac{\partial N}{\partial t} = \cos\theta \frac{\partial}{\partial x} \left(K_s N \frac{\partial N}{\partial x} \right) + \sin\theta \frac{\partial}{\partial x} (K_s N), \quad (2)$$

where f is the effective porosity of the groundwater ($\text{m}^3 \text{m}^{-3}$). As implied by the Boussinesq equation, the groundwater properties are considered homogeneous across tiles, including K_s , f , and the bedrock slope θ . $\frac{\partial N}{\partial x}$ denotes the groundwater hydraulic gradient between adjacent tiles. At each tile, the water table (H) is determined by the balance of the soil–groundwater and the lateral groundwater fluxes (Eq. 3).

$$H_i^{j+1} = H_i^j + \left(\frac{r_i^j - q_{li}^j}{\rho f} \right) \Delta t \quad (i = 1, \dots, n_{\text{HB}}), \quad (3)$$

where H_i^j is the hydraulic head of the water table (m) at the i th height band (HB) at the j th time step. Based on the continuity equation, the $N - H$ relationship can be established as $H = \frac{N}{\cos\theta}$. n_{HB} is the total number of HBs in the catchment. ρ is the liquid water density (1000 kg m^{-3}), and r_i^j is the liquid water flux between the soil–bedrock column and

the water table (mms^{-1}) at the i th height band at the j th time step. The physical time step Δt of the model is 1800 s (30 min). q_{1i}^j denotes the divergence of the lateral groundwater flux at the i th height band at the j th time step and is set to the difference between the lateral divergence from/to the HBs immediately above and below. In particular, in HB₁ (i.e., the height band nearest to the river reach), q_{11}^j is determined by the balance of the groundwater discharge from HB₂ and the baseflow (if going up the reach) or the channel infiltration (if going down the reach). If multiple tiles exist in an HB, r_i is effectively calculated by arithmetically averaging the r values from each tile belonging to the i th height band.

LM4-SHARC also resolves the reach-scale streamflow dynamics, and the resulting hydrograph is used as the time-dependent boundary condition at the interface between the stream and the hillslope. The Saint-Venant continuity equation with a kinematic wave assumption is solved for the streamflow dynamics. For a kinematic wave, the momentum equation assumes that the energy grade line is parallel to the streambed (De St Venant, 1871; Strelkoff, 1970). Thus, the friction slope (S_f) is assumed to be equal to the streambed slope (S_o), physically implying that the flow is primarily governed by gravity and friction (Eq. 4). Manning’s equation was used to express the flow resistance implied in Eq. (4). Using Manning’s equation, the stream discharge Q ($\text{m}^3 \text{s}^{-1}$) can be related to the cross-section area of flow U (m^2) using a coefficient that depends on channel roughness, slope, and geometry (α) and an exponent determined by the flow characteristics (β) (Eq. 5) (Manning, 1891). Given that at any cross-section, Q and U are functionally related as Eq. (5), the continuity equation can be expressed as Eq. (6).

$$S_f = S_o \tag{4}$$

$$U = \alpha Q^\beta \tag{5}$$

$$\frac{dQ}{dy} = \frac{\partial Q}{\partial y} + \frac{\partial Q}{\partial t} \frac{dQ}{dU} = \frac{2q_1 H_1}{\rho}, \tag{6}$$

where $\frac{dQ}{dU}$ is the kinematic wave celerity (ms^{-1}), and y denotes the river flow direction coordinate. H_1 is the vertical thickness of the groundwater at HB₁, which effectively defines the wetted perimeter at the stream–hillslope interface (m). $\frac{2q_1 H_1}{\rho}$ is the lateral inflow per unit channel width ($\text{m}^2 \text{s}^{-1}$), as q_1 is estimated from Eq. (1). Once the spatial derivative of the stream discharge Q (i.e., $\frac{dQ}{dy}$) is resolved by Eq. (6), the reach outflow (i.e., discharge at the catchment outlet) was used to inversely estimate the river stage, which was in turn used to determine the lateral hydraulic gradients between the river stage and the water table in HB₁.

We changed the soil column’s bottom boundary condition from a zero-flux to a variable-flux boundary condition to allow for the two-way interaction between the soil columns and the newly introduced groundwater domain. As shown in Fig. 1, due to the hydraulic gradient between the bottommost soil layer and the water table, the vertical liquid flux (r in

Eq. 3) defines the soil column bottom boundary condition if the water table is deeper than 10 m from the land’s surface (Eqs. 7 and 8). However, if the water table is within 10 m from the ground, the vertical liquid flux at the soil base is considered zero since the equipotential line is assumed to be vertical in the saturated zone, so the vertical hydraulic gradient is zero (Eq. 7). The variable-flux boundary condition chosen allows the soil bottom drainage (SBD) at 10 m in depth. This enables the consideration of the effects of groundwater on the unsaturated soil processes depending on the groundwater configuration, such as groundwater properties and the water table depth.

$$r_i^j = \begin{cases} \frac{K_{\text{stog},i}^j (\psi_{\text{btm},i}^j + \Delta L_i^j)}{\Delta L_i^j} & (\text{water table depth} > 10 \text{ m}) \\ 0 & (\text{water table depth} \leq 10 \text{ m}) \end{cases} \tag{7}$$

$$\Delta L_i^j = e_{\text{btm},i} - (H_i^j + hl_i \tan \theta), \tag{8}$$

where r_i^j is the vertical liquid water flux (mms^{-1}), and $K_{\text{stog},i}^j$ is the hydraulic conductivity between the bottommost unsaturated soil layer and the water table at the i th height band at the j th time step (mms^{-1}), which is calculated by the harmonic mean of hydraulic conductivity values in the bottommost soil layer and the groundwater. $\psi_{\text{btm},i}^j$ is the soil matrix potential at the i th height band at the j th time step (m) and $e_{\text{btm},i}$ the elevation of the central node in the bottommost soil layer (m). $w_{t_i}^j$ is the pressure head of the water table at the i th height band at the j th time step (m). hl_i is the total hillslope length from the reach of the i th height band characterized by the HMC approach (m). Therefore, Eq. (8) shows that ΔL_i^j is the distance between the bottommost soil node and the water table at the i th height band at the j th time step (m) when considering the bedrock slope θ .

2.3.2 Two-way energy transport and conservation in LM4-SHARC

LM4-SHARC accounts for the phase change of water in the groundwater domain by calculating the ice content according to the groundwater temperature (Eq. 9).

$$ws_{\text{gw},i}^{j+1} = \begin{cases} ws_{\text{gw},i}^j - \min \left(ws_{\text{gw},i}^j (T_{\text{gw},i}^j - T_{\text{freeze}}) \frac{hc_{\text{gw},i}^j}{hf} \right) & (ws_{\text{gw},i}^j > 0, T_{\text{gw},i}^j > T_{\text{freeze}}) \\ ws_{\text{gw},i}^j + \min \left(wl_{\text{gw},i}^j (T_{\text{freeze}} - T_{\text{gw},i}^j) \frac{hc_{\text{gw},i}^j}{hf} \right) & (wl_{\text{gw},i}^j > 0, T_{\text{gw},i}^j < T_{\text{freeze}}), \end{cases} \tag{9}$$

where wl_{gw} and ws_{gw} are the liquid and ice contents of the groundwater, respectively (–). T_{gw} is the groundwater temperature (K) and T_{freeze} the freezing point of 273.15 K. hf is the latent heat of fusion, a constant of $3.3358 \times 10^5 \text{ J kg}^{-1}$. hc_{gw} is the dynamic heat capacity of the groundwater ($\text{JK}^{-1} \text{ m}^{-2}$). The groundwater temperature in each HB

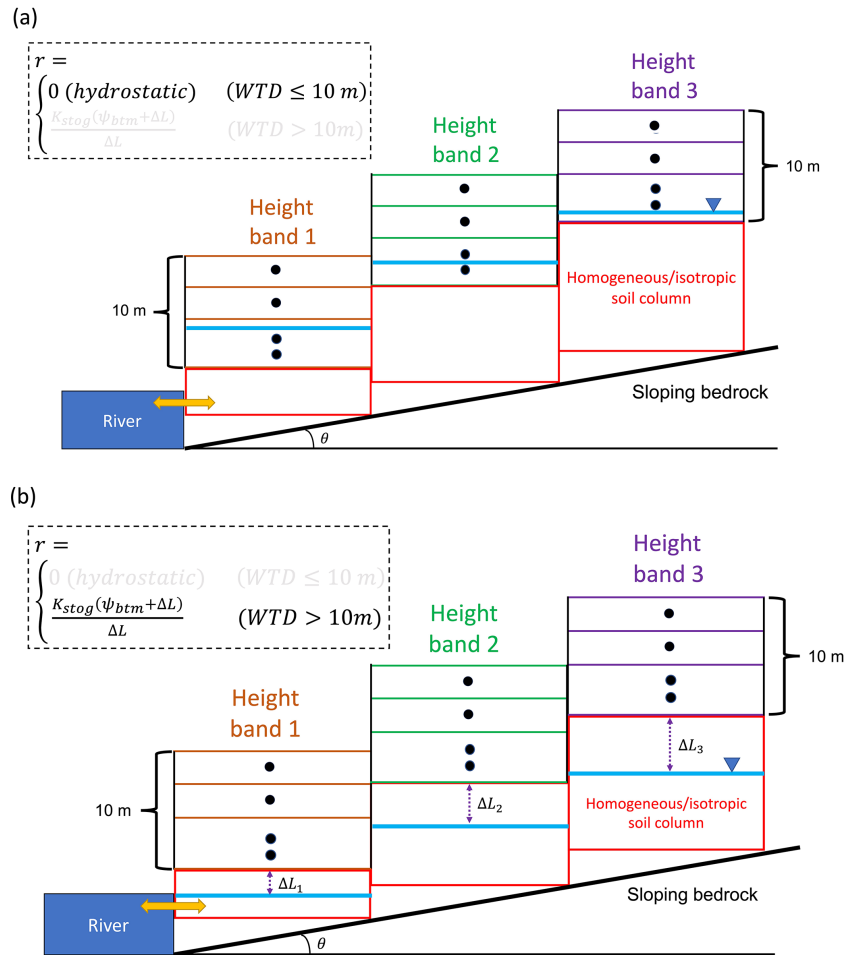


Figure 1. The LM4-SHARC’s modified boundary condition (BC) at the soil base depending on the depth of the water table. **(a)** The equipotential line is assumed to be vertical (i.e., hydrostatic) if the depth of the water table is less than 10 m; **(b)** the soil base BC changes from zero flux to variable flux according to the hydraulic gradient between the bottommost soil layer and the water table. WTD denotes the water table depth. The black dots refers to the nodes located at the center of each soil layer.

is dynamically updated, taking into account time-dependent heat capacity (hc_{gw}) and heat fluxes conducted and advected from/to adjacent flow domains. Similar to the water-table-dependent boundary condition at the soil base, the heat flux boundary condition at the soil base is also affected by the groundwater condition since the water advection is zero if the water table depth is less than 10 m (Eq. 10).

$$\delta_{v_i}^j = \begin{cases} (\delta_{v_{adv,i}}^j + \delta_{v_{cnd,i}}^j) & (\text{water table depth} > 10 \text{ m}) \\ \delta_{v_{cnd,i}}^j & (\text{water table depth} \leq 10 \text{ m}), \end{cases} \quad (10)$$

where $\delta_{v_{cnd,i}}^j$ is the vertical heat conduction flux, and $\delta_{v_{adv,i}}^j$ is the advected heat flux between the soil column and groundwater ($\text{J m}^{-2} \text{ s}^{-1}$). The equations of the vertical heat conduction ($\delta_{v_{cnd,i}}^j$) and advection ($\delta_{v_{adv,i}}^j$) based on the dynamic groundwater heat capacity are addressed in Appendix A. The direction of $\delta_{v_{adv,i}}^j$ is determined by the water flux (i.e., downward recharge or upward capillary) according to the

hydraulic gradient. Thus, the soil columns and groundwater temperatures are simulated with the modified heat flux boundary condition (at the soil column base) from constant-thermal to variable-thermal fluxes in LM4-SHARC. The lateral heat transport in the groundwater domain ($\delta_{l_i}^j$) is another component that determines the groundwater temperature (Eqs. 11–13).

$$\delta_{1i}^j = -q_{1,i}^j \text{clw} \left(T_{\text{gw},i}^j - T_{\text{freeze}} \right) \quad (i = n_{\text{HB}}) \quad (11)$$

$$\delta_{1i}^j = \text{clw} \left[q_{1,i+1}^j \left(T_{\text{gw},i+1}^j - T_{\text{freeze}} \right) - q_{1,i}^j \left(T_{\text{gw},i}^j - T_{\text{freeze}} \right) \right] \quad (12)$$

$(i = 2, \dots, n_{\text{HB}} - 1)$

$$\delta_{11}^j = \begin{cases} \text{clw} \left[q_{1,2}^j \left(T_{\text{gw},2}^j - T_{\text{freeze}} \right) - q_{1,1}^j \left(T_{\text{gw},1}^j - T_{\text{freeze}} \right) \right] & (q_1 \geq 0) \\ \text{clw} \left[q_{1,2}^j \left(T_{\text{gw},2}^j - T_{\text{freeze}} \right) - q_{1,1}^j \left(T_{\text{stream}}^j - T_{\text{freeze}} \right) \right] & (q_1 < 0), \end{cases} \quad (13)$$

where δ_{1i}^j is the lateral groundwater heat flux ($\text{J m}^{-2} \text{s}^{-1}$). The groundwater temperature at the i th height band at the j th time step ($T_{\text{gw},i}^j$) is determined based on the time-dependent heat capacity and the vertical/lateral heat fluxes from/to the i th height band at the j th time step (Eq. 14).

$$T_{\text{gw},i}^{j+1} = T_{\text{gw},i}^j - \frac{\left[\left(\delta_{\text{v},\text{end},i}^j - \delta_{\text{v},\text{adv},i}^j \right) + \delta_{1i}^j \right] \Delta t}{\text{hc}_{\text{gw},i}^j} \quad (14)$$

For the lateral heat exchange fluxes between the stream and HB_1 , the stream temperature is considered if the channel loses water to the riparian zone (i.e., $q_1 < 0$). The stream temperature is estimated considering how much heat flows into the stream from the hillslopes and out of it through the catchment outlet (Eq. 15).

$$T_{\text{stream}}^{j+1} = T_{\text{stream}}^j + \frac{\left[\left(\delta_{11}^j + \delta_{\text{unsat}_i}^j \right) - \text{clw} \left(T_{\text{stream}}^j - T_{\text{freeze}} \right) \frac{\rho Q_i^j}{A^j} \right] \Delta t}{\text{hc}_{\text{stream}}^j}, \quad (15)$$

where $\delta_{\text{unsat}_i}^j$ is the heat flux advection from the unsaturated soil to the reach by interflow ($\text{J m}^{-2} \text{s}^{-1}$). Q_i^j is the stream discharge at the catchment outlet ($\text{m}^3 \text{s}^{-1}$), and A^j is the flow area at the outlet at the i th height band at the j th time step. $\text{hc}_{\text{stream},i}^j$ is the heat capacity of the catchment outflow according to the streamflow hydrograph at the outlet ($\text{J m}^{-2} \text{K}^{-1}$). Consequently, in LM4-SHARC, states and fluxes for each domain are determined by accounting for the two-way water–heat exchanges among the unsaturated soil, the hillslope aquifer, and the river.

2.4 The streamflow recession analysis for the groundwater parameterization

Brutsaert and Nieber (1977) showed that the time derivative of the recession hydrograph can be expressed as a function of the streamflow Q (Eq. 16). Since the analytical solutions to the Boussinesq equation can be recast in the form

of a power law, the Boussinesq groundwater can be effectively characterized based on groundwater parameters such as K_s , f , the initial saturated groundwater thickness N_{ini} , and the contributing area A (Hong and Mohanty, 2023a, b; Rupp and Selker, 2006; Brutsaert and Nieber, 1977; Szilagyi et al., 1998; Tashie et al., 2021).

$$-\frac{dQ}{dt} = a Q^b, \quad (16)$$

where b is a constant, and a is a function of the groundwater properties. Since the geometric similarity of unit-width Boussinesq groundwater throughout the entire catchment is assumed, the catchment outflow Q was estimated as $Q = 2q_1 H L_{\text{reach}}$ (where L_{reach} is the length of the reach). This geometric similarity assumption is also reflected in the numerical estimation of the catchment outflow hydrograph (Eq. 6). Because the recession parameters a and b are readily estimated by a logarithmic regression of Q on $-\frac{dQ}{dt}$, streamflow observations can be used to infer the effective groundwater properties.

2.4.1 Selecting analytical models

Theoretical catchment outflow from the Boussinesq groundwater yields two hydraulic regimes: the early (i.e., high-flow) and late (i.e., low-flow) time domains. Since LM4-SHARC considers sloping groundwater, we only consider the analytical solution of the Boussinesq equation for sloping groundwater. We used the analytical solutions obtained by Brutsaert (1994), considering their applicability for a broader range of bedrock slopes (Pauritsch et al., 2015). Thus, the recession parameter a for the early time domain (a_{early}) is expressed in Eq. (17), with b_{early} set to 3.0, and the parameter a for the late time domain (a_{late}) is defined in Eq. (18), with b_{late} set to 1.0. As a result, Eqs. (17) and (18) were used to interpret the intercept (i.e., $\log(a)$) and slope (i.e., b) of the logarithmic regression of Q on $-\frac{dQ}{dt}$ derived from observational streamflow data in the early time and late time domains, respectively.

$$a_{\text{early}} = \frac{1.33}{K_s f N_{\text{ini}}^3 L^2 \cos \theta}, \quad b_{\text{early}} = 3.0 \quad (17)$$

$$a_{\text{late}} = \frac{\pi^2 p K_s N_{\text{ini}} L^2}{f A^2} \cos \theta \left[1 + \left(\frac{B}{T_{\text{ini}}} \tan \theta \right)^2 \right], \quad (18)$$

$$b_{\text{late}} = 1.0,$$

where A is the subsurface drainage area (m^2) that effectively contributes to the recession slope characteristics, and p is a constant set to 1/3. B is the contributing groundwater's characteristic length (m) under the geometric similarity assumption, calculated as $B = A/(2L)$.

2.4.2 Event-scale recession analysis

This study recognized that the recession parameters from the point cloud data (i.e., collective recession data) could be artifacts of the variability in individual recession events (REs) (Jachens et al., 2020a; Tashie et al., 2020; Karlsen et al., 2019). To fill this gap, we performed an event-scale recession analysis to account for the variability in recession slope characteristics among individual REs, which results in different estimates of groundwater properties. The onset of the REs was started 5 d after the peak to exclude the influence of overland flow (i.e., runoff) on the streamflow hydrograph. We examined the continuous decline in daily discharge observational data to decide the duration of an RE, and the end date of an RE was determined when the daily stream discharge was at its lowest.

For each RE, we performed the logarithmic regression of Q on $-\frac{dQ}{dt}$ in bi-logarithmic space. The time derivative of Q ($\frac{dQ}{dt}$) was computed based on the daily streamflow. Each RE's hydraulic regime transition point, from the early to the late time domain, was identified daily using the method suggested by Hong and Mohanty (2023a). In this method, the transition point from the early time to the late time domain is determined based on the abrupt and most noticeable change in R^2 values from linear regression with a fixed slope of 3.0 while incrementing the number of the data pair of $\log(-\frac{dQ}{dt})$ – $\log(Q)$ in descending order in the bi-logarithmic space. To this end, we selectively used the daily streamflow time series of REs that lasted for more than 20 d to get enough data pairs of $\log(Q)$ – $\log(-\frac{dQ}{dt})$ to distinguish the hydraulic regime. For further details on the method for identifying the hydraulic regime transition in an event-scale recession analysis, see Hong and Mohanty (2023a). The process workflow of our method determining the catchment-scale hydraulic diffusivity parameters by the combined use of the analytical and numerical models is described in Fig. 2.

2.5 Experimental design for model comparison

LM4-SHARC's new parameterization for the groundwater and its interaction with the soil and river was evaluated based on a comparison of the baseflow and near-surface soil moisture/temperature outputs from the retrospective runs of LM4-SHARC and LM4-HydroBlocks, respectively. For the spinup, we used periodically cycled GSWP3 10-year forcing. The GSWP3 forcing from 1901–1910 was used repeatedly in each model cycle until a steady-state in the groundwater-related model variables was reached by replacing the initial condition for a new spinup with the final output state from the previous cycle. The groundwater-related variables include (1) soil moisture content (SMC) in the bottommost soil layer, (2) baseflow, and (3) the water table, and we considered the model to have reached a steady state if the simulated difference between the end of the n th and $(n - 1)$ th cycle for the variables used satisfied our criteria simultaneously. In this

study, we set the criteria for each variable to $0.001\text{m}^3\text{m}^{-3}$ (0.1 %) for the bottommost layer's SMC, 0.1md^{-1} for the baseflow, and 0.001 m for the water table. We note that the soil–groundwater two-way fluxes (r in Eq. 7) were additionally considered in LM4-SHARC to evaluate a steady state, as they are a new variable in LM4-SHARC.

Using the confirmed steady state outputs as an initial condition, both model configurations were compared for the period from 1 October 2003 to 30 September 2014 (WYs 2004–2014, 11 years). In this study, the in situ precipitation and air temperature were assimilated into the GSWP3 forcing data by being directly inserted (i.e., direct insertion data assimilation). This was done to improve the accuracy of the model outputs, given the inconsistency between the GSWP3 forcing and the local meteorological conditions, which was primarily due to scale differences (i.e., $0.5^\circ \times 0.5^\circ$ vs. 1km^2). We considered precipitation and air temperature to be the most significant atmospheric variables determining the catchment's water and energy budgets, so the in situ precipitation and air temperature data obtained during the evaluation period replaced the corresponding variables in the GSWP3 forcing. LM4-HydroBlocks and LM4-SHARC were then operated using the identical forcing inputs. Two statistical metrics, Pearson's correlation coefficient (R^2) and RMSE (root-mean-square error), were used to evaluate the temporal agreement of the modeled hydrologic outputs vs. the corresponding observations and errors, respectively.

3 Study area and observational data

3.1 Study area and the HMC parameters

The study area is the 1km^2 Providence Creek P301 headwater catchment in the Sierra National Forest, Nevada (Fig. 3a). The P301 headwater catchment is one of eight primary headwater catchments of the Kings River Experimental Watershed (KREW) project. The US Department of Agriculture's (USDA) Pacific Southwest Research Station initiated and operated the KREW project as part of the National Science Foundation's Southern Sierra Critical Zone Observatory (Jepsen et al., 2016; Hunsaker et al., 2012). The eight catchments are clustered into two groups, Providence and Bull, and, in this study, the P301 headwater catchment that belongs to Providence Creek was selected. We selected the P301 catchment because it contains only one first-order reach with no tributaries flowing into the reach. Since the connectivity between the catchment and reach had to be established to develop LM4-SHARC, the hydrologic configuration of the P301 catchment was considered ideal for evaluating the newly developed model.

Surface elevation in the P301 catchment range from 1755 to 2114 m (Hunsaker et al., 2012) on an average topographic slope of 19° . The length of the first-order reach in the P301 catchment is 1.5 km, and the average width of the reach

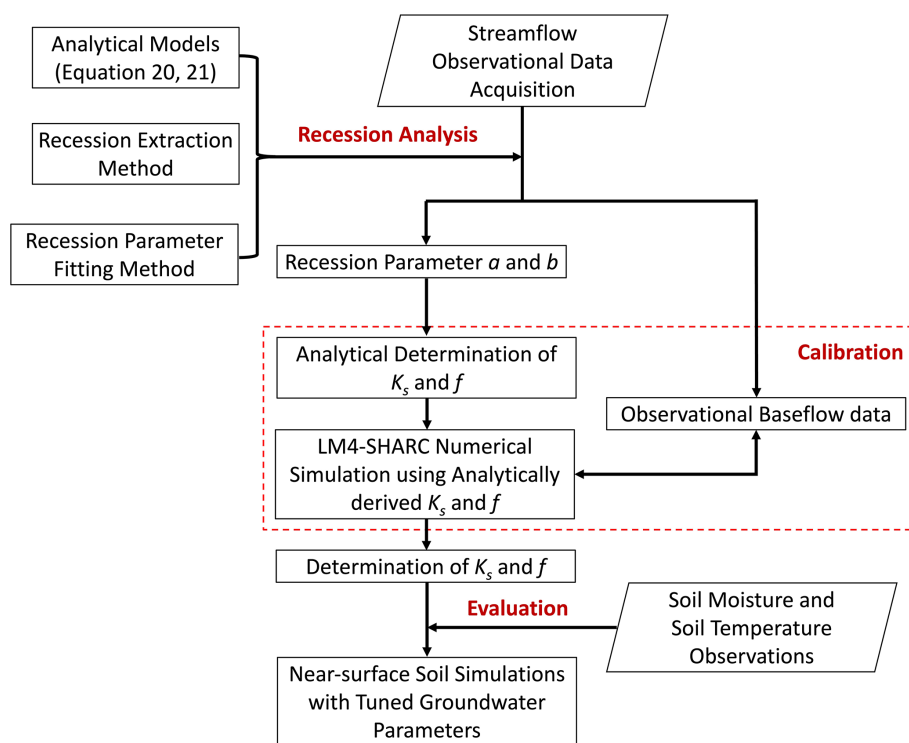


Figure 2. A workflow diagram of the method presented to calibrate and evaluate the catchment-scale groundwater diffusivity parameters based on streamflow recession analysis and the combined use of the analytical models and LM4-SHARC numerical simulations.

was approximated at 10 m to define the channel geometry. The Providence P301 catchment represents a rain–snow mixed-conifer forest site with an annual mean precipitation of 1315 mm yr^{-1} . The site has a Mediterranean climate, with cool, wet winters and dry summers from approximately May through October (Safeeq and Hunsaker, 2016). During the WY 2004–2014 evaluation period, about 90 % of precipitation occurred between November and June. The mean annual air temperature was measured at $6.8 \text{ }^\circ\text{C}$. Precipitation falls as a mix of rain and snow, and precipitation transitions from mostly rainfall to mostly snow at approximately 2000 m in elevation (Bales et al., 2018; Hunsaker et al., 2012).

In compiling the land input dataset for the given catchment, k was set to 1, as the study area is a single catchment. The surface elevation data from 3DEP DEM were used as the sole variable to account for the intra-catchment variability in terrain properties, and each HB’s spatial extent was determined using a dh value of 20 m. p was set to 1; however, we note that the number of tiles in each HB can increase (or decrease) depending on factors such as natural mortality, land use, and fire events applied to each tile. The stream flowline was also delineated according to the 3DEP DEM using an area threshold of $100\,000 \text{ m}^2$. Consequently, the P301 catchment was clustered into six HBs with a delineated area of 0.9904 km^2 and a stream length of 1.3 km, nearly identical to the field measurements (Fig. 3b–d).

3.2 Observations for models evaluation

3.2.1 Streamflow and baseflow

Streamflow observations measured at the outlet of the P301 catchment were used. The primary stream height measurement device is an ISCO 730 air bubbler (Teledyne Isco). Backup stage measurements were initially obtained using an AquaRod capacitance water level sensor (Advanced Measurements and Controls, Inc.) or a Telog pressure transducer (Trimble Water, Inc.). The stream height is measured at 15 min intervals and converted into the discharge rate using the standard rating curve supplied by the flume and weir manufacturers (Bales et al., 2018). The stream discharge monitoring began in September 2003 (i.e., WY 2004). The streamflow was averaged daily from WY 2004 to WY 2014 (i.e., 1 October 2003–30 September 2014, 11 years in total) and was used to evaluate the daily basis simulation outputs in this study. We derived the daily baseflow rate from the daily observational stream flow data using the baseflow separation method suggested by Szilagyi and Parlange (1998). The baseflow separation method applied assumes that the drainage from the Boussinesq groundwater maintains the stream recession flow. Thus, the baseflow separation method ensures consistency between the analytical models applied and the observational baseflow data. Szilagyi and Parlange’s separation method requires the catchment area, length of

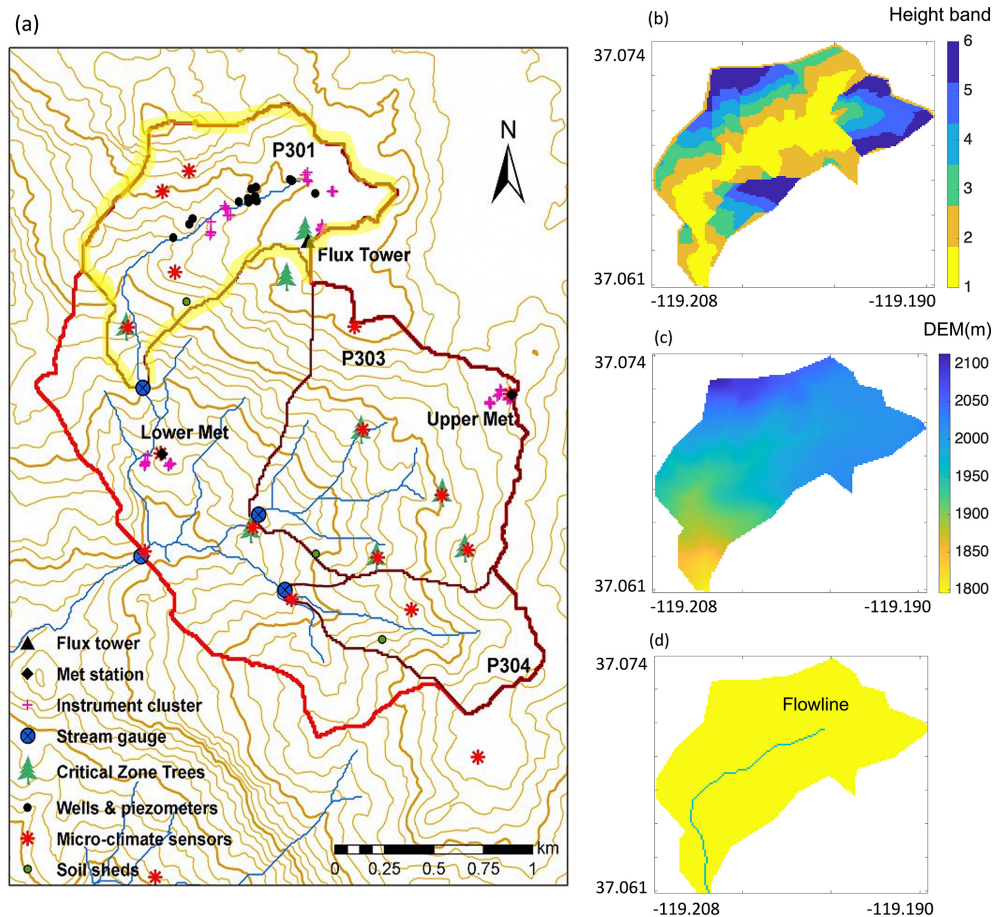


Figure 3. (a) The yellow highlighted area indicates the spatial extent of the headwater catchment P301. While the soil moisture and temperature data measured at the Lower Met station are not within the P301 catchment, the measurements from the Lower Met station were used due to the station's proximity to the catchment and the superior quantity and continuity of the observational data compared to other measurement points within the P301 catchment. (b) Using the HMC method, the P301 catchment was clustered into six different height bands (HBs). (c) The digital elevation map (DEM) of the P301 and (d) the flowline delineated by the land input dataset preprocessing. Further division of the study catchment into hillslopes is indicated in Fig. S1 in the Supplement.

channels, and the initial groundwater thickness as parameters. For consistency, the values of these parameters were determined following the parameters used in the analytical models for the recession analysis (Eqs. 17 and 18). Therefore, the catchment area (A) and channel length (L) are considered static according to the field measurement (area – 1 km^2 ; length – 1.5 km). Applying the A value of 1 km^2 assumes that the entire catchment area contributes to the streamflow recession characteristics. However, if only a portion of the catchment contributes to baseflow, this assumption could create uncertainty in the baseflow estimates (Hong and Mohanty, 2023a). The initial saturated groundwater thickness N_{ini} , assumed to be constant at a value of 10 m across the individual REs and equivalent to the initial value of groundwater thickness applied in LM4-SHARC, could also be a source of uncertainty in the baseflow estimates. However, due to the lack of observation-based A and N_{ini} values, this

study conducted baseflow separation under these assumptions for the two parameters.

3.2.2 Soil moisture, soil temperature, and snow depth

The soil volumetric water content (i.e., soil moisture content, SMC, and soil temperature, ST) were measured at the Lower Met using ECHO-TM sensors (METER Group). The SMC and ST were measured at 10, 30, 60, and 90 cm below the soil surface, and the measurements were used as representative values for each soil depth above and below each sensor (Bales et al., 2018). The measurements at the Lower Met station were used due to its proximity to the P301 headwater catchment ($\sim 480 \text{ m}$ to the outlet of the P301 catchment). The elevation difference between the Lower Met station and the nearest drainage point (i.e., reach) is about 13 m . At the station, the sensor nodes were installed at locations with different canopy coverage characteristics, such as drip

edge, under canopy, and open canopy, to account for the effect of shade (i.e., radiation interception due to the canopy) on SMC and ST. Snow depth was also measured at the same Met station, and the distance to snow (or soil if no snow cover exists) was measured using an acoustic depth sensor located 3 m above the soil surface (Judd Communication LLC) (Bales et al., 2018). The sensors were installed in 2008 (i.e., WY 2009); however, due to the availability of the 10 cm depth SMC observations at the open canopy spot, the observations and simulations were compared from WY 2009 to WY 2012 (4 years). The SMC and ST observations measured at the depth nearest to the land surface (10 cm) were used to evaluate the near-surface modeled outputs from LM4-SHARC and LM4-HydroBlocks.

3.2.3 Meteorological observations

The meteorological data for model forcing were obtained concurrently from a weather station located at the Lower Met station (elevation of 1750 m). We used the SMC, ST, and snow observations. Precipitation was measured with a Belfort 5-780 shielded weighing rain gauge (Belfort Instrument) located 3 m above the soil surface. The air temperature sensor is 6 m above the soil surface (Vaisala Corporation) (Bales et al., 2018).

4 Results and discussion

4.1 Suitability of the select analytical model

We compared the analytical and numerical solutions of baseflow flux for different combinations of the groundwater diffusivity (D) and bedrock slope (θ) to test the agreement required for the application of analytically derived groundwater properties to the numerical groundwater domain. We defined the case of slow groundwater flow (SLW), with K_s of $2 \times 10^{-3} \text{ mm s}^{-1}$ and f of 0.2; K_s of $9 \times 10^{-3} \text{ mm s}^{-1}$ and f of 0.05 for normal groundwater flow (NRM); and K_s of $2 \times 10^{-2} \text{ mm s}^{-1}$ and f of 0.02 for fast groundwater flow (FST). In each diffusivity case, the baseflow was also simulated for distinct bedrock slope conditions, such as relatively flat bedrock with a $\tan\theta$ of 0.001° (MLD) and steep bedrock with a $\tan\theta$ of 0.2° (STP). The hydraulic conditions, except for D and $\tan\theta$, for this gaining reach experiment are as follows: (1) river stages at j th time step at the discharge boundary follow a power function using the initial river stage $h_s^j = h_s^0 \times t^{-0.01}$, which represents a falling limb of the hydrograph after peak discharge. (2) The initial head difference between the initial saturated groundwater thickness (N_{ini}) and h_s was set to the half of N_{ini} (i.e., $h_s^j|_{j=0} = N_{\text{ini}}/2$). Identical hydraulic conditions were applied to both analytical and numerical simulations.

The temporal agreement and the total magnitude of groundwater divergence fluxes per unit width (i.e., q_1N) were investigated during the recession duration of 15 d. As shown

in Fig. 4, q_1N decreases as the hydraulic gradient between the river stage h_s and water table (at height band 1) decreases due to the discharging groundwater. For the representativeness of the time series, the R^2 and RMSE were estimated for the average q_{lat} at each time step (i.e., $qN_{\text{ave}} = \frac{qN_{\text{stp}} + qN_{\text{mld}}}{2}$). The RMSE and R^2 were calculated at $0.00088 \text{ m}^2 \text{ s}^{-1}$ and 0.98 in the SLW case, $0.0015 \text{ m}^2 \text{ s}^{-1}$ and 0.97 in the NRM case, and $0.0027 \text{ m}^2 \text{ s}^{-1}$ and 0.96 in the FST case. Although the gaps between the analytical and numerical qN_{ave} showed a bit of a drop in the agreement as the groundwater diffusivity increased, we understand that the similarly good estimation of the daily cumulative numerical qN_{ave} and temporal agreement (R^2 0.96–0.98) could compensate for the gap and yield groundwater discharge estimates that are accurate enough to model the streamflow recession. Specifically, care needs to be taken when the analytical model is used to tune the numerical Boussinesq groundwater with extremely high D and steeper bedrock (Fig. 4c). Except for such hydraulically extreme cases that are unrealistic in real catchments (i.e., $D > 1.0 \text{ mm s}^{-1}$), the numerical simulation of the Boussinesq groundwater's discharge with analytically tuned parameters can be justifiable.

4.2 Event-scale recession analysis and calibration

4.2.1 Quantifying the uncertainty in hydraulic diffusivity estimates

We performed the event-scale recession analysis to estimate the effective groundwater properties of the study catchment. Following the extraction criteria (Sect. 2.4.2), 18 individual REs were extracted from the 13-year series of streamflow observations, and the recession parameters a and b were estimated for each hydraulic regime (Table 1). Figure 5a presents an example showing how an individual RE was analyzed with the selected analytical models for its early and late time domains using daily streamflow data from the P301 catchment. Once the recession parameters a and b were estimated for each RE, the range of bedrock slope θ should be adequately constrained in order to determine the hydraulic diffusivity parameters K_s and f . The characteristic length B is calculated by assuming the geometric similarity between both sides of the hillslope in the catchment; thus, $B = A/(2L)$. The following criteria were applied to define the upper and lower bounds of θ : (1) the effective porosity f should range from 0.1 % to 20.0 % (Brutsaert and Lopez, 1998; Tashie et al., 2021; Troch et al., 1993; Hong and Mohanty, 2023b, a; Heath, 2004), and (2) the catchment-scale effective groundwater lateral hydraulic conductivity K_s cannot exceed 1.0 mm s^{-1} (Fan and Miguez-Macho, 2011; Gómez-Hernández and Gorelick, 1989; Tashie et al., 2021).

The variability in recession characteristics was quantified by the recession parameter a (since b is fixed) in the early and late time domains (i.e., a_{early} and a_{late}) (Fig. 5b). Essentially, this parameter provides insights into the variabil-

Table 1. Recession period, recession characteristics, and parameters a and b for each recession event under different diffusivity conditions. The variability in the parameter a indicates the variability in the distinct diffusivity of groundwater across individual recession events.

RE	Recession period (duration in days)	$\log(a_{\text{early}})$	$\log(b_{\text{early}})$	$\log(a_{\text{late}})$	$\log(b_{\text{late}})$
1	06.04.2004–12.05.2004 (37 d)	−3.98	3.0	−6.95	1.0
2	02.07.2004–28.07.2004 (27 d)	−3.72		−6.76	
3	20.05.2005–08.06.2005 (20 d)	−3.61		−7.10	
4	18.06.2005–14.08.2005 (58 d)	−3.55		−6.75	
5	23.05.2006–12.06.2006 (21 d)	−2.71		−6.84	
6	14.06.2006–17.07.2006 (34 d)	−2.99		−6.33	
7	30.07.2006–06.09.2006 (39 d)	−3.01		−7.66	
8	05.05.2007–30.05.2007 (26 d)	−2.75		−5.72	
9	07.06.2007–07.07.2007 (31 d)	−2.12		−6.76	
10	05.06.2008–12.07.2008 (38 d)	−1.88		−6.45	
11	14.07.2008–18.08.2008 (36 d)	−2.75		−5.98	
12	03.05.2009–29.05.2009 (27 d)	−2.12		−6.60	
13	01.07.2009–05.08.2009 (36 d)	−2.84		−5.61	
14	06.06.2010–01.08.2010 (57 d)	−1.92		−6.67	
15	03.08.2010–27.08.2010 (25 d)	−1.99		−6.11	
16	01.08.2011–22.08.2011 (22 d)	−2.86		−6.32	
17	27.04.2012–24.05.2012 (28 d)	−2.33		−7.12	
18	22.05.2014–10.06.2014 (20 d)	−2.70		−5.49	

ity in the groundwater's effective properties dependent on the memory effects of the catchment (e.g., groundwater storage). The 98 % confidence intervals for a_{early} and a_{late} were estimated to be $[-3.10, -2.43]$ and $[-6.81, -6.20]$, respectively, from the analysis of 18 REs. Following the above criteria, the value distributions of K_s , f , and θ corresponding to the ranges of a_{early} and a_{late} were estimated. The uncertainty in K_s , f , and θ was then quantified by determining the intersection range between K_s , f , and θ values derived from the lowest a (i.e., $\log(-\frac{dQ}{dt}) = -3.10\log(dQ) - 6.81$) and the highest a (i.e., $\log(-\frac{dQ}{dt}) = -2.43\log(dQ) - 6.20$). The upper and lower bounds of K_s were thus estimated as $[0.0026, 0.0138 \text{ mm s}^{-1}]$. For f , the bounds were identified as $[0.033, 0.190]$, and for θ they were $[5.0, 17.0^\circ]$. Also, for each set of data pairs of $\log(-\frac{dQ}{dt}) - \log(dQ)$ with the lowest and highest a , the relationship between hydraulic diffusivity (per unit of wetted perimeter) D and the bedrock slope θ follows a power function. Consequently, we considered the (θ, D) space in which solutions of D and θ (potentially representing the catchment average behavior) could exist to be further constrained by the range of parameters specified, K_s , f , and θ , between the power functions (Fig. 5b).

4.2.2 Calibrating groundwater properties based on baseflow flux accuracy

While the groundwater properties vary across the REs due to the catchment's memory effect, the groundwater properties that represent the long-term average behavior of the groundwater need to be tuned in the LM4-SHARC groundwater domain. Based on the value range of θ and D that was identified, we tried to determine a single pair of (θ, D) that

shows the optimal accuracy by comparing the modeled and observed baseflow data. The modeled baseflow fluxes were estimated by summing the liquid fluxes from saturated soil to the stream, and baseflow observations during this study's evaluation period (from 1 October 2003 to 30 September 2014, WYs 2004–2014) were used for calibration. We considered a (θ, D) that best represents the temporal dynamics and magnitudes of baseflow observations to be the calibrated (θ, D) for the study catchment.

We identified the fact that a specific parameter pair (θ, D) can be specified in the uncertainty space. Figure 6 shows that (1) the temporal agreement between the modeled and observed baseflow was generally related to the magnitude predictions and (2) the most significant improvements to R^2 and RMSE (from LM4-HydroBlocks to LM4-SHARC) were identified at the same time at a specific pair of (θ, D) . We found the most pronounced improvements to the baseflow predictions in LM4-SHARC compared to LM4-HydroBlocks, with an R^2 improvement of 0.155 and an RMSE reduction of 0.220 mm d^{-1} at $\theta = 5.0^\circ$ and $D = 4.6 \times 10^{-2} \text{ mm s}^{-1}$. The R^2 and RMSE of the LM4-SHARC-derived baseflow estimates vs. the observations (over the 11 years) were estimated at 0.402 and 0.556 mm d^{-1} , while those of the LM4-derived baseflow estimates were estimated at 0.247 and 0.776 mm d^{-1} , respectively. With the calibrated groundwater properties, we confirmed that the recession behaviors of the P301 catchment streamflow hydrograph over the 11-year evaluation period were generally better captured in the LM4-SHARC baseflow estimates compared to those of LM4-HydroBlocks (Fig. 7).

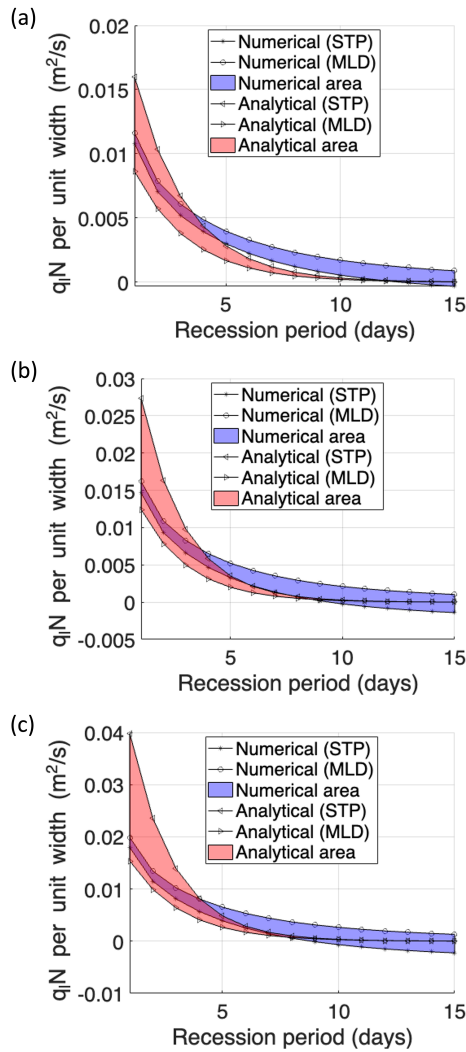


Figure 4. Comparison between the baseflow flux solutions derived from the analytical and numerical models selected. (a) The SLW case with K_s of $2.0 \times 10^{-3} \text{ mm s}^{-1}$ and f of 0.20. (b) The NRM case with K_s of $9.0 \times 10^{-3} \text{ mm s}^{-1}$ and f of 0.05. (c) The FST case with K_s of $2.0 \times 10^{-2} \text{ mm s}^{-1}$ and f of 0.02. The numerical and analytical areas refer to the uncertainty bands of numerical and analytical solutions due to slope variations.

4.3 The effect of groundwater on near-surface water and energy balances

4.3.1 Near-surface soil moisture and temperature predictions

Using the calibrated groundwater diffusivity properties, we assessed the effects of groundwater-induced soil processes on the near-surface water and energy budgets. We applied $\theta = 5.0^\circ$ and $D = 4.6 \times 10^{-2} \text{ mm s}^{-1}$ to tune the Boussinesq groundwater domain (in LM4-SHARC) and compared the soil moisture and temperature estimates at 10 cm in depth from both configurations. The evaluation of near-surface

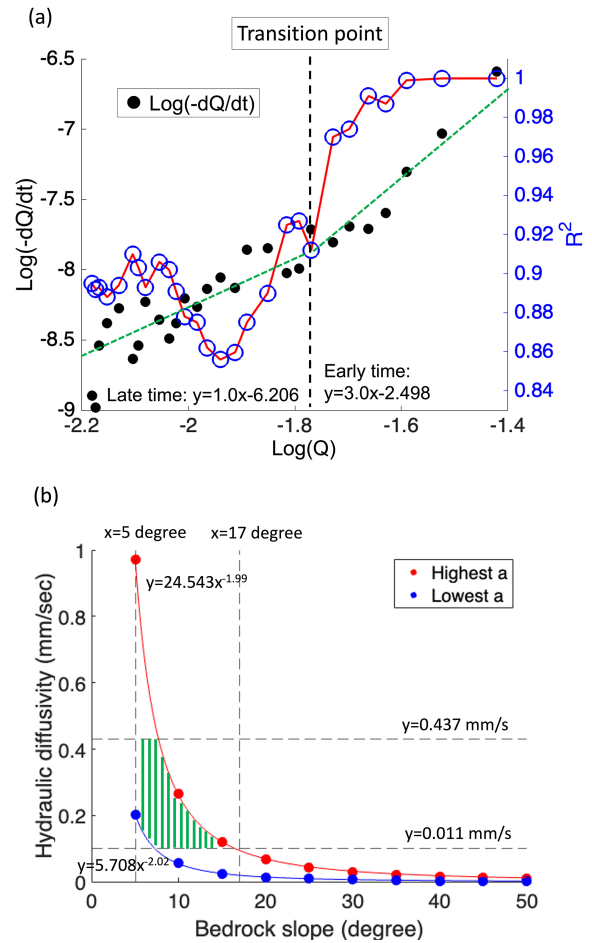


Figure 5. (a) An example showing how the transition point of a hydraulic regime (from the early to the late time domain) is determined from an individual recession event. Understanding the combined recession parameters a_{early} and a_{late} enables us to infer the groundwater properties such as hydraulic diffusivity D . (b) The uncertainty in the groundwater properties D and θ was constrained by the power function relationship between them, which results from the variability in the recession parameter a across individual recession events.

soil moisture and temperature predictions was performed for 4 years, from WY 2009 to WY 2012. Figure 8a–d show the comparative time series of 10 cm depth soil moisture, soil temperature, and snow mass among LM4-HydroBlocks, LM4-SHARC, and in situ observations in each WY. From all 4 years of the evaluation period, we identified the fact that soil bottom drainage (SBD) facilitation due to the modified boundary condition from zero-flux to variable-flux BC at the base of the soil columns (Fig. 1a) significantly affected the soil moisture content (SMC) at 10 cm in depth. The time series of 10 cm SMC from LM4-SHARC showed significantly reduced SMC compared to that of LM4-HydroBlocks. The entire soil column reached full saturation too quickly in LM4-HydroBlocks, and the downward liquid transport fa-

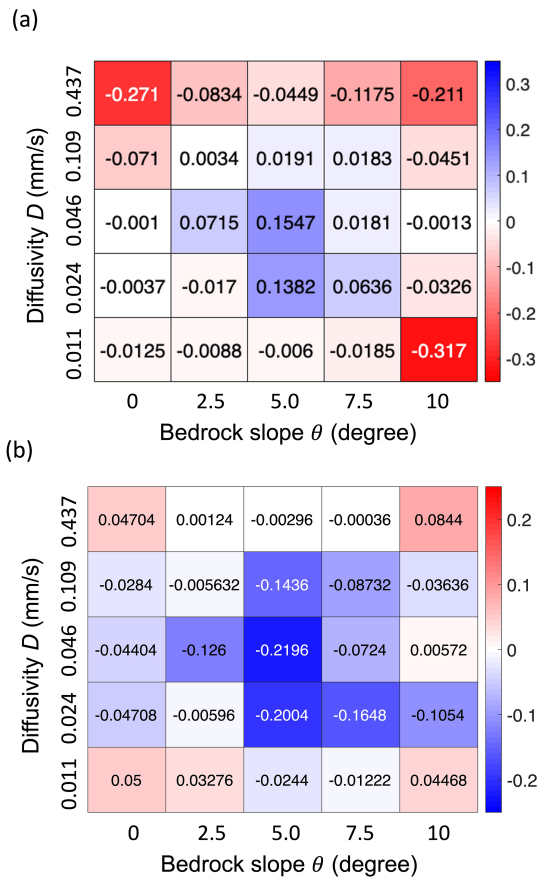


Figure 6. A pair of groundwater parameters (θ, D) was specified to calibrate the catchment groundwater. **(a)** The R^2 difference (improvement) between LM4-SHARC and LM4-HydroBlocks ($R^2_{\text{LM4-SHARC}} - R^2_{\text{LM4-HydroBlocks}}$). Here, R^2 denotes the coefficient of determination between the modeled baseflow and observations for the 11-year evaluation period. **(b)** For the same period, the RMSE difference (reduction) between LM4-SHARC and LM4-HydroBlocks ($\text{RMSE}_{\text{LM4-SHARC}} - \text{RMSE}_{\text{LM4-HydroBlocks}}$) was evaluated.

cilitated due to drier lower layers of soil in LM4-SHARC was found to be effective in correcting the soil's wet biases. As a result, the temporal variability and total storage of the observed SMC was better captured in LM4-SHARC. The average of the four R^2 values from the yearly time series comparison (i.e., 2009–2012) improved from 0.831 (LM4-HydroBlocks) to 0.849 (LM4-SHARC), while noticeably reducing the average RMSE from $0.0722 \text{ m}^3 \text{ m}^{-3}$ (LM4-HydroBlocks) to $0.0425 \text{ m}^3 \text{ m}^{-3}$ (Fig. 9a and b).

We also found that the enhanced representation of SMC resulted in better capture of near-surface soil temperature dynamics. The decreased SMC reduces the evapotranspiration rate, especially during daytime, leading to increased sensible heat in the soil's energy balance. Also, the reduced soil heat capacity due to decreased SMC (i.e., liquid water) increased the soil temperature under the given net radi-

ation. Consequently, we identified the fact that soil temperature predictions at the 10 cm depth showed significant improvements in LM4-SHARC, primarily during warmer seasons, and showed better agreement with the observations. The soil temperature estimates from both model configurations showed similar values when the surface was covered by snow (i.e., snow depth > 0). In this case, the soil is insulated by snow and thus variations in soil water predicted by the two model configurations do not lead to appreciable differences in soil temperature. From the improved skill in predicting near-surface soil temperature in LM4-SHARC, we conclude that the soil columns with applied zero-flux BC at 10 m in depth could hold too much soil water due to the shallow water table imposed. This overestimation in soil water content could lead to an inaccurate description of the land energy balance (e.g., overestimated ET/less sensible heat) and thus to biases in soil temperature. We note that the average of the four R^2 values from the yearly time series comparison between the modeled and observed soil temperature increased from 0.952 to 0.957, and the RMSE significantly decreased, from 2.77 to 1.67 K (Fig. 9c and d).

The simulated snow depth from both LM4-HydroBlocks and LM4-SHARC generally showed reasonable agreement with the measured snow depth in the catchment. We note that the meltout timing of snow involves the mutual influence of soil temperature. Specifically, we noticed that the timing of snow meltout is affected by soil temperature, as the warmer ground expedites the melting. The meltout timing of snow in the 4 evaluation years was represented sooner in LM4-SHARC than LM4-HydroBlocks due to increased soil temperature (Fig. 8). Also, as the snow melted out, the soil temperature quickly increased, as the soil was no longer insulated by snow, leading to a higher correlation with surface air temperature.

4.3.2 Sensitivity of soil water storage to stream–groundwater diffusivity

After examining the enhancement in the catchment-scale water and energy balance, we further explored to what extent groundwater properties affect soil processes in LM4-SHARC, focusing on soil water storage (SWS). We investigated the sensitivity of SWS to the groundwater properties θ and D by estimating differences in SWS per unit area (i.e., ΔSWS (kg m^{-2})) between LM4-SHARC and LM4-HydroBlocks. ΔSWS was calculated by subtracting the total SWS per unit area of the soil columns (in the catchment), which was derived from LM4-HydroBlocks, from that of LM4-SHARC at the end of the simulation. To this end, we need to verify that the model reaches a steady state for the given groundwater properties. The variability in SWS (ΔSWS) is evaluated only after the steady state has been reached. Figure 10a–d show that the simulations of the four variables reduce their deviations (from the previous cycle) as the cycles progress and that the model was considered to

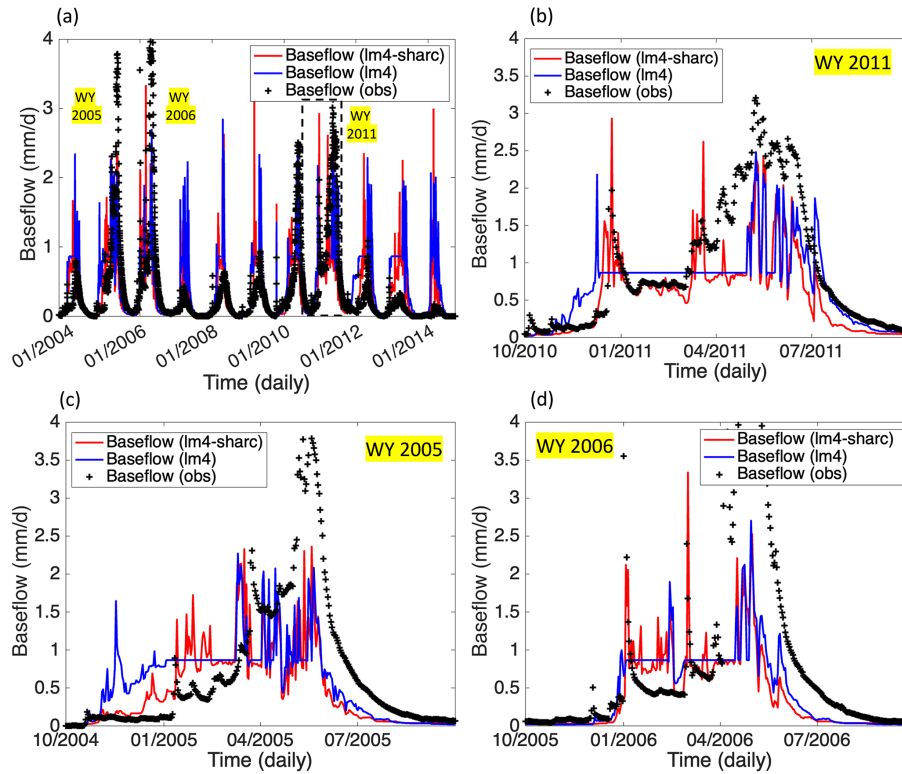


Figure 7. Comparative time series of baseflow estimates from LM4-SHARC (red), LM4-HydroBlocks (LM4 in the legend, blue), and the corresponding observations. (a) Daily time series of the baseflow data over the evaluation period of 11 years. (b) Time series in water year (WY) 2011, (c) time series in water year (WY) 2005, and (d) time series in water year (WY) 2006. The streamflow recession behavior was generally better represented in LM4-SHARC compared to LM4-HydroBlocks.

be steady state based on the agreement of variables from the n th and $(n - 1)$ th cycles. We also identified the fact that the groundwater properties influenced the time to reach a steady state (i.e., spinup time), as the variations in groundwater diffusivity affect the flux velocity. More detail can be found in Fig. S2 in the Supplement.

Δ SWS due to groundwater discharge was investigated based on the steady state of the model. The working hypothesis here is that the lateral groundwater discharge could facilitate the SBD by inducing downward hydraulic gradients from the subsurface soil and water table. Figure 11a shows the variability and magnitude of the entire catchment's annual mean of Δ SWS (over 10 years) according to θ and D (y -axis values denote the average Δ SWS per soil column). It is noticeable that the annual mean of Δ SWS gradually increased negatively (i.e., less SWS in LM4-SHARC) with increasing D until the water table dynamics did not significantly affect hydraulic gradients between subsurface soil and the water table (i.e., lack of groundwater storage). This happened when the downward groundwater recharge fluxes were continuously less than q_1 in the groundwater domain. Also, the catchment's total Δ SWS was found to be lower if the slope is steep. While the steeper bedrock showed a greater q_1 with the increased gravitationally driven discharge fluxes,

Δ SWS values were more noticeably affected by the lowered water table due to relatively flatter bedrock in the groundwater domain. In the case of this catchment, it was also observed that for D values greater than 0.1 mm s^{-1} , the lack of groundwater storage occurred irrespective of the bedrock slope.

Moreover, SBD facilitation due to the groundwater lateral discharge was more noticeable, as the height band was farther from the reach. The annual mean of Δ SWS per unit area gradually decreased from HB_6 ($\sim 120 \text{ kg m}^{-2} \text{ yr}^{-1}$) to HB_1 ($\sim 10 \text{ kg m}^{-2} \text{ yr}^{-1}$), with a sharp decrease at HB_1 (Fig. 11b). It was also noticeable that the distinct D made less difference among the values of Δ SWS per unit area, as the height band was farther from the reach. Figure 11c and d illustrate that the effects of groundwater flow conditions on the SWS variability could be more significant in the riparian/river valley compared to the hilltop area. This also implies that the groundwater effects on the water content in the partially saturated soil depend on the depth of the water table, leading to higher sensitivity of SMC to groundwater diffusivity if the water table is shallow than if it is deep. While the HB_6 annual Δ SWS per unit area (m^2) values were the greatest among HB_1 to HB_6 , ranging from $121.45 \text{ kg m}^{-2} \text{ yr}^{-1}$ (VSW, very slow) to 121.89 mm (FST), the Δ SWS difference (between VSW and FST) of $0.44 \text{ kg m}^{-2} \text{ yr}^{-1}$ was minor compared to the result

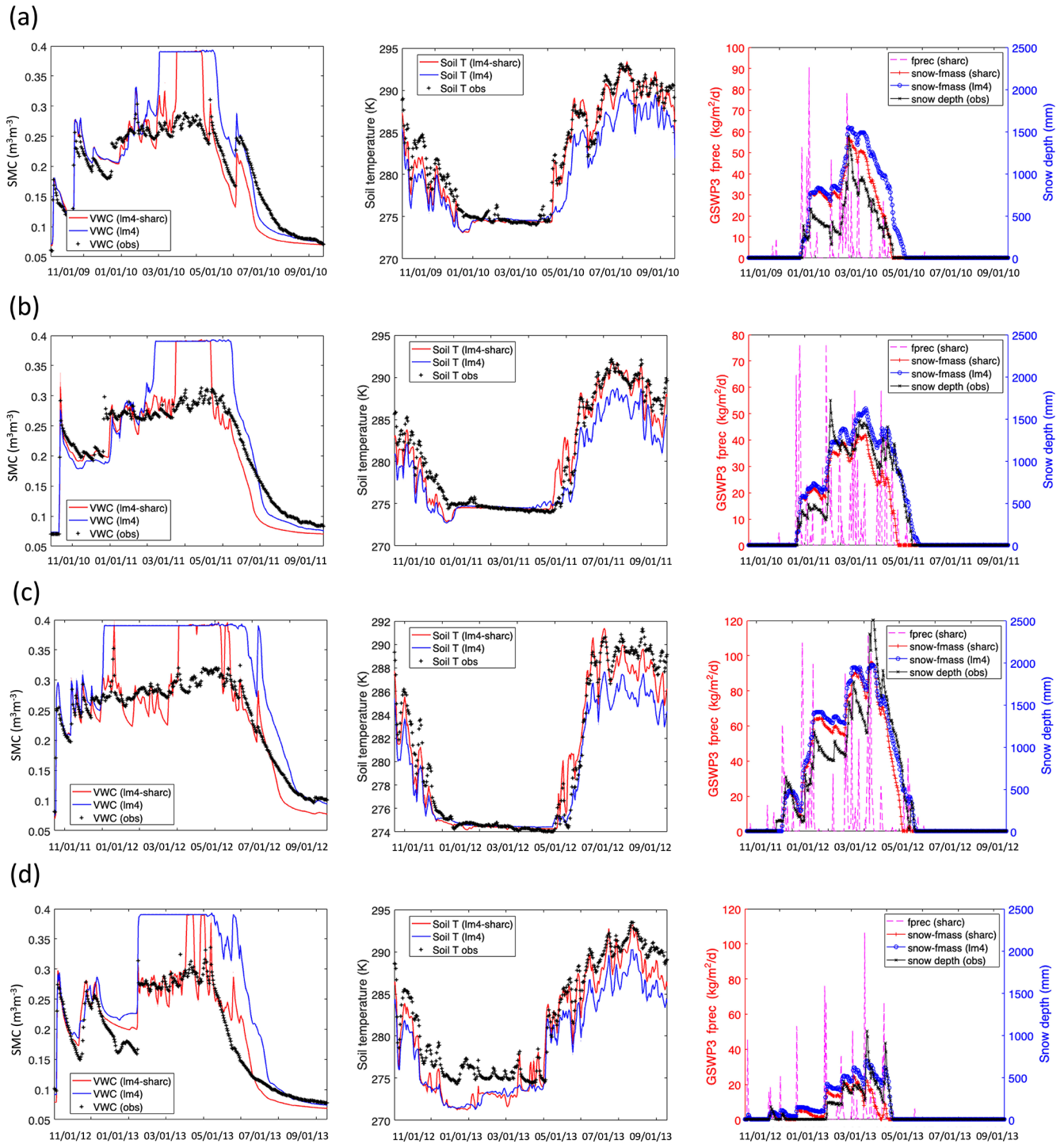


Figure 8. Comparative yearly time series of soil moisture, temperature, and snow mass for the evaluation periods in (a) WY 2009, (b) WY 2010, (c) WY 2011, and (d) WY 2012. For all years, higher soil temperature at 10 cm in depth due to drier soil was generally identified. GSWP3 fprec (y-axis label in the third column) stands for frozen precipitation, denoting the snowfall rate in units of $\text{kg m}^{-2} \text{d}^{-1}$.

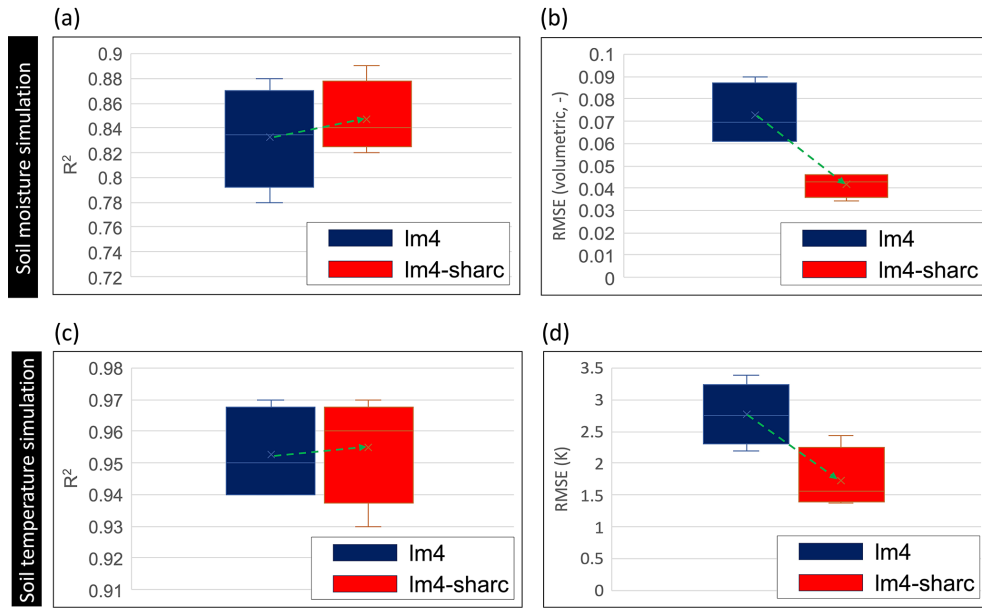


Figure 9. (a, b) The average of the four R^2 from the yearly time series comparison of 10 cm depth soil moisture (i.e., 2009–2012) improved from 0.831 (LM4-HydroBlocks) to 0.849 (LM4-SHARC), while significantly reducing the average RMSE from $0.0722 \text{ m}^3 \text{ m}^{-3}$ (LM4-HydroBlocks) to $0.0425 \text{ m}^3 \text{ m}^{-3}$. (c, d) R^2 from the yearly time series comparison between the modeled and observed soil temperature at 10 cm in depth increased from 0.952 to 0.957, and the RMSE significantly decreased, from 2.77 to 1.67 K. LM4 denotes LM4-HydroBlocks.

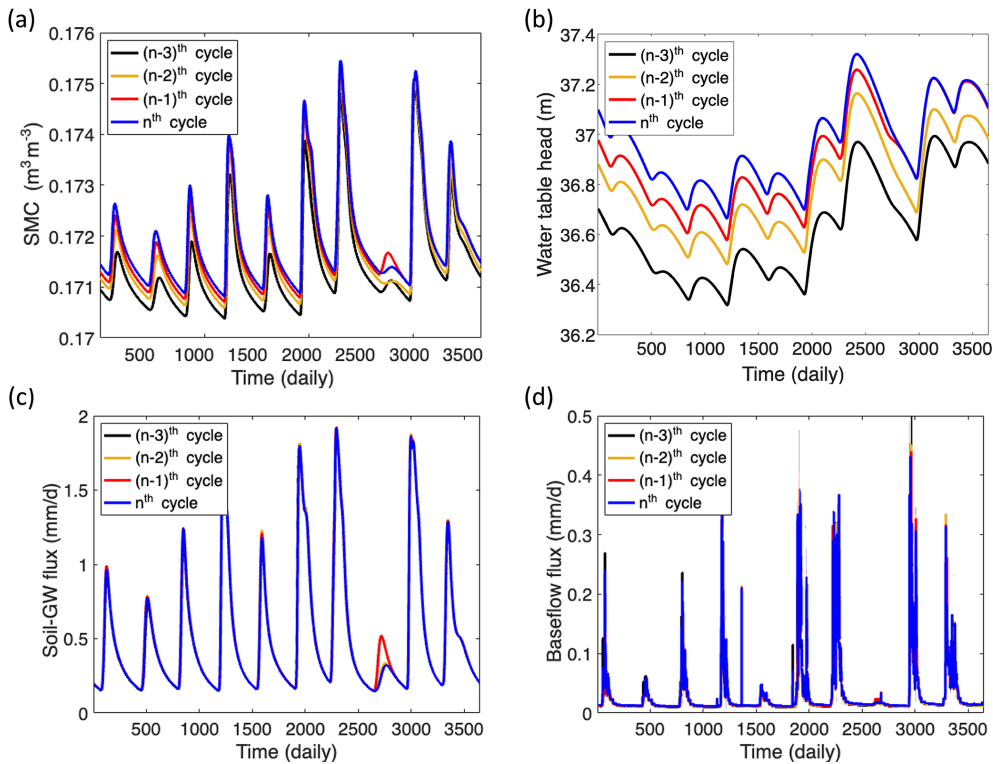


Figure 10. (a) The volumetric soil moisture content at the bottommost soil layer. (b) The groundwater table. (c) The groundwater recharge fluxes (positive). (d) The baseflow fluxes reduce their deviations (compared to the previous cycle) as the cycles progress. They are considered steady state based on the differences between the n th and $(n - 1)$ th cycle.

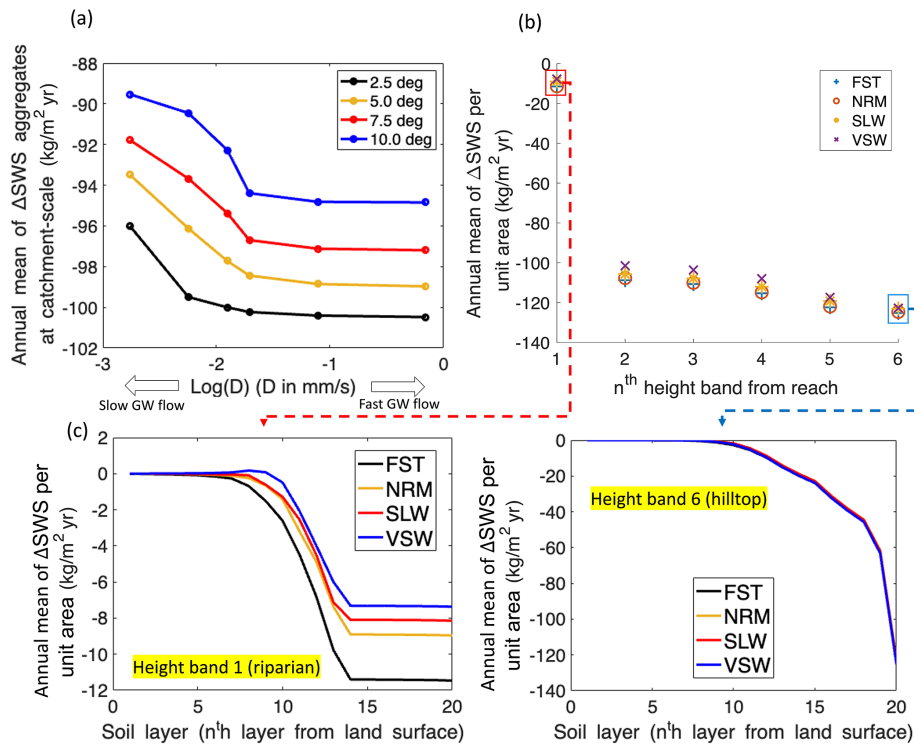


Figure 11. (a) Sensitivity of soil water storage (SWS) to the diffusivity D and bedrock slope θ . (b) The soil drainage (i.e., downward recharge) due to the groundwater lateral discharge was more noticeable as the height band was farther from the reach. (c, d) More-significant effects of groundwater flow conditions on the SWS in the riparian zone compared to those in the hilltop area.

from HB_1 . From HB_1 , the Δ SWS values were found to range from $7.75 \text{ kg m}^{-2} \text{ yr}^{-1}$ (VSW) to $11.70 \text{ kg m}^{-2} \text{ yr}^{-1}$ (FST), yielding a difference of $3.94 \text{ kg m}^{-2} \text{ yr}^{-1}$ due to the groundwater diffusivity.

4.4 Distinct vegetation characteristics in the catchment due to hydrologic contrast

The water convergence due to the groundwater lateral flow induces the hydrologic contrast at the catchment scale, leading to distinct vegetation characteristics depending on the distance from the river (Fan et al., 2019). Here, we used the modeled leaf-area index (LAI) to infer distinct vegetation characteristics (i.e., plant density) in the study catchment. The LAI was simulated from 1901 to 2014 (114 years) in both model configurations without spinup using the GSWP3 forcing with assimilated in situ precipitation/air temperature data (i.e., GSWP3 (January 1901–September 2003) and in situ data (October 2003–September 2014)). Comparison of the LAI time series between LM4-SHARC and LM4-HydroBlocks revealed that the differences in the LAI in the hilltop area (i.e., HB_6) were more dramatic than those in the riparian zone (i.e., HB_1) as the vegetation evolved (Fig. 12a and b). With a closer look at the LAI time series for the recent 4 water years (i.e., WYs 2009–2012, as studied in Sect. 5.3.1) in the riparian zone and on the hilltop, we found

that the LAI contrast was, moreover, more significant during the warmer season in the hilltop area, while the overall trend in the LM4-SHARC-derived LAI evolved comparably to that of the LM4-derived LAI in the riparian zone (Fig. 12c and d). The hydrologic convergence in the riparian zone explains the comparable LAI dynamics in the riparian zone, as the subsoil domain readily saturated by the converging water impedes SBD, leading to higher water retention in the partially saturated soil. Different soil moisture availability (between HB_1 and HB_6) resulting from the contrasting SBD dynamics is thus emphasized by the density of plants, especially during the warm(er) season when the plants yield higher transpiration rates. The varied transpiration rates in LM4-SHARC also consistently explain the LAI contrast due to different soil moisture conditions. We found that the transpiration rate on the hilltop was reduced by 29.5 % in LM4-SHARC, while the rate was reduced by 10.3 % in the riparian zone in LM4-SHARC (Fig. 12e and f). Overall, we found that the variations in SWS, transpiration, and LAI simulations are consistent in that the groundwater convergence to the river valley intensified the catchment's contrasting hydrologic states.

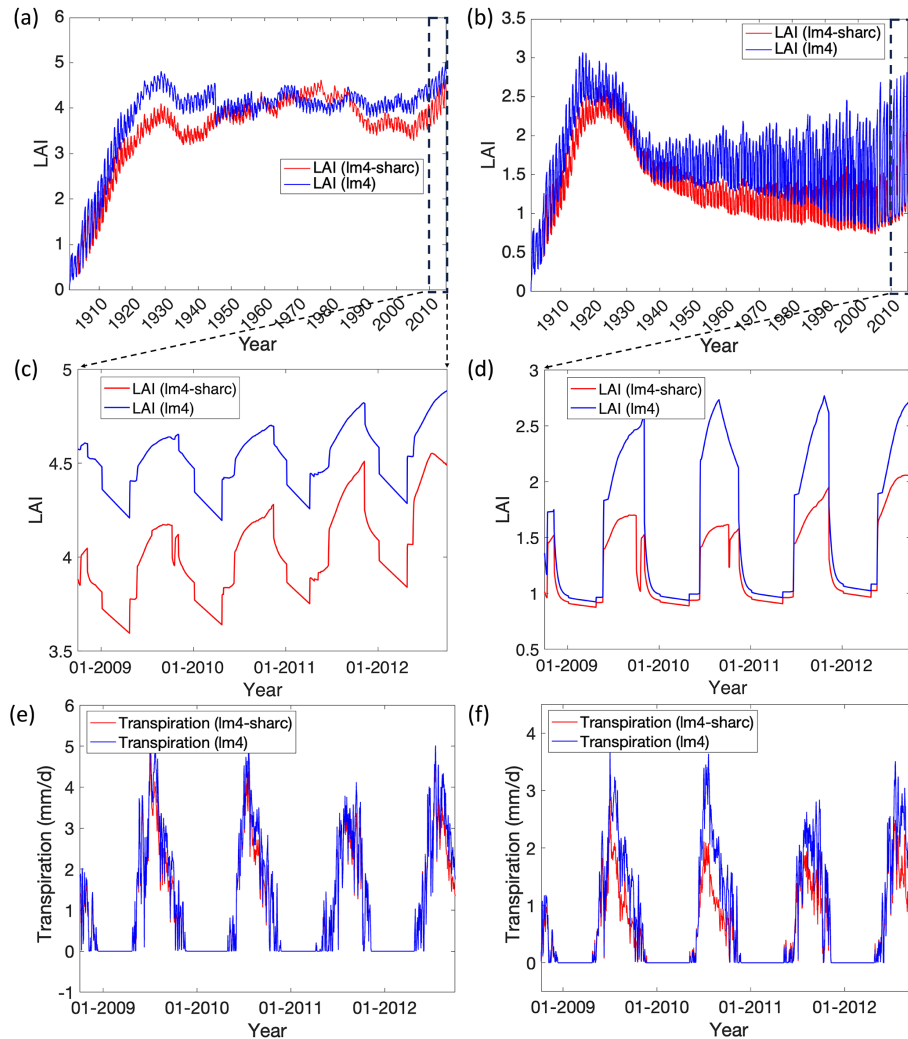


Figure 12. Temporal evolution of LAI from 1901 to 2014 (114 years) using GSWP3 forcing at (a) the riparian zone (HB1) and (b) the hilltop (HB6). The differences in the LAI (between LM4-SHARC and LM4-HydroBlocks) in the hilltop area were more dramatic than those in the riparian zone. (c, d) The LAI time series for the recent 4 WYs (i.e., WY 2009–2012) and (e, f) the time series of the plant transpiration rate (mm d^{-1}) for the same period. LM4 denotes LM4-HydroBlocks. The effects of varying transpiration dynamics in LM4-SHARC on evapotranspiration estimates are discussed in Fig. S3 in the Supplement.

4.5 Applicability of LM4-SHARC in an ESM

4.5.1 Testing LM4-SHARC in various climatic and orographic zones

To support the implementation of LM4-SHARC in the GFDL ESM, we need to investigate the performance of LM4-SHARC in various climatic and orographic zones. For this purpose, three additional headwater catchments were selected based on their precipitation and topographic slope characteristics (Fig. 13a): the Musselshell (MT), Maine (ME), and Clearwater (WA) headwater catchments. The precipitation and slope characteristics of these sites vary from those of the Clearwater catchment (the wettest, 3136 mm yr^{-1} , and steepest, 0.547 m m^{-1} , catchment) to

those of the Musselshell catchment (the driest, 397 mm yr^{-1} , and flattest, 0.094 m m^{-1}). In this experiment, the groundwater properties in the additional catchments were assumed to be identical to those of the P301 headwater catchment, so diffusivity D was set to 0.046 mm s^{-1} and the slope of the groundwater bedrock θ was assumed to be 5° in all three headwater catchments. The steady state of the groundwater and any adjacent flow domains was also ensured, and the evaluation was performed using the 10-year model outputs.

We tried to identify whether the model is robust under diverse conditions by examining the consistency between hydrologic characteristics and indices. The hydrologic indices include (1) the runoff ratio (i.e., the ratio of streamflow to precipitation) and (2) the baseflow coefficient (i.e., the ratio of baseflow to precipitation).

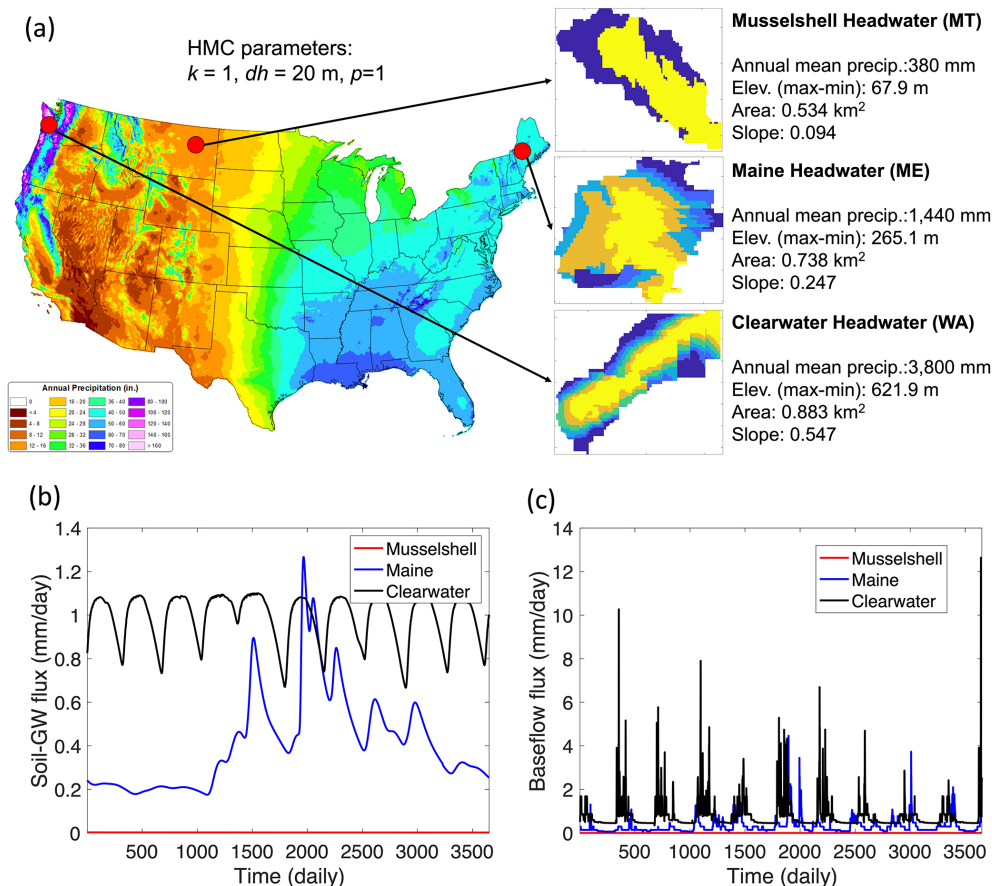


Figure 13. (a) The three additional sites are marked (red dots) on the PRISM 30-year normal precipitation map (2022, PRISM Group, Oregon State University). The sites include the Musselshell headwater (MT), Maine headwater (ME), and Clearwater headwater (WA) catchments. (b) LM4-SHARC's daily soil-groundwater exchange fluxes (mm d^{-1}) over 10 years from the three catchments and (c) the simulated daily baseflow fluxes (mm d^{-1}) for the corresponding period and sites.

The annual mean baseflow/streamflow of each catchment was estimated at $10.9/49.8 \text{ mm yr}^{-1}$, $105/364.0 \text{ mm yr}^{-1}$, and $281/1897.4 \text{ mm yr}^{-1}$ from the Musselshell, Maine, and Clearwater catchments, respectively. Using the respective annual mean precipitation of each catchment (i.e., Musselshell – 397.2 mm yr^{-1} ; Maine – $1223.1 \text{ mm yr}^{-1}$; and Clearwater – $3136.6 \text{ mm yr}^{-1}$), each catchment's baseflow coefficient/runoff ratio was estimated at $0.028/0.125$, $0.086/0.298$, and $0.230/0.604$. The established positive correlation between the baseflow coefficient and runoff ratio is consistent with what has been reported in previous studies (Cheng et al., 2021; Ouyang et al., 2018) (Table 2). We also found that the value gradients of baseflow/recharge estimates correspond to what is expected from the slope/precipitation gradients. For example, the highest yield of baseflow from the Clearwater catchment (i.e., 0.77 mm d^{-1}) can be explained by its high precipitation and steep slope, which contribute to higher drought flow (i.e., baseflow during dry seasons) and peaks. Also, the partially saturated soil found in most parts of the Musselshell catchment explains the mini-

mal baseflow amounts due to the lack of groundwater storage from the limited groundwater recharge (Fig. 13b and c).

4.5.2 Inferring groundwater properties for global-scale SHARC simulations

Since the parameterization scheme SHARC presented in this paper relies on the observationally derived recession characteristics of streamflow, a method must be developed to quantify the recession variability (using the parameter a) of catchments with no streamflow information. This necessity is particularly emphasized, as the SHARC scheme will ultimately be used for global simulations. While not providing specific research findings, this section aims to discuss several possible approaches based on existing studies. Essentially, we expect that existing global-scale datasets of soil, topography, climatology, and remotely sensed hydrologic data can be used in conjunction to infer the ungauged catchments' recession characteristics. For example, the significant correlation between the recession parameter a and catchments' soil/geology attributes (from a global database)

Table 2. The hydrologic characteristics and indices of the three additional headwater catchments from the simulated outputs with the GSWP3 forcing from 1901 to 1910. The annual mean values of each catchment are estimated from the 10-year cycle based on the confirmed steady state.

		Headwater catchment		
		Musselshell (MT)	Maine (ME)	Clearwater (WA)
Hydrologic characteristics	Annual mean precipitation	397.2 mm yr ⁻¹	1223.1 mm yr ⁻¹	3136.6 mm yr ⁻¹
	Annual mean streamflow	49.8 mm yr ⁻¹	364 mm yr ⁻¹	1894.7 mm yr ⁻¹
	Annual mean baseflow	10.9 mm yr ⁻¹	105 mm yr ⁻¹	281 mm yr ⁻¹
	Annual mean groundwater recharge	0.21 mm yr ⁻¹	150.4 mm yr ⁻¹	360.2 mm yr ⁻¹
Indices	Runoff ratio	0.125	0.298	0.604
	Baseflow coefficient	0.028	0.086	0.230

was established by means of simple regression analysis or machine learning (ML) techniques (e.g., a random forest) (Zecharias and Brutsaert, 1988; Hong and Mohanty, 2023b; Tashie et al., 2021; Cai et al., 2021). In particular, Hong and Mohanty (2023b) examined the relationship between the local catchment's mean soil permeability derived from the STATSGO2 soil map dataset, and the recession parameter a , enabling parameterization of Boussinesq groundwater at a large-scale that encompassed three major basins in Texas. In this context, we will explore the applicability of existing global soil datasets such as SoilGrids (Poggio et al., 2021) and GSDE (Shangguan et al., 2014) by identifying generalizable relationship(s) between the catchment average soil properties and recession characteristics (established among gauged catchments globally) to tune LM4-SHARC's groundwater parameters at ungauged catchments. For example, Hong and Mohanty (2023b) found a significant relationship between the STATSGO2-derived catchment average soil permeability and observationally derived recession parameters, and they parameterized the catchment-scale groundwater in all 33 436 catchments (including 40 gauged and 33 396 ungauged catchments). Furthermore, in addition to global soil property data, the utility of remotely sensed and global in situ hydrologic data will be studied to understand whether the groundwater recession behavior could be predicted using machine learning (ML) frameworks (e.g., random forests). ML frameworks that efficiently compute the correlation between the target (e.g., model parameters) and independent variables (e.g., observationally derived data) greatly facilitate parameter inference in data-scarce or ungauged regions, improving the scalability of processes. We aim to explore globally available remote sensing hydrologic/ecologic data, such as SMAP, GRACE-FO, MODIS, SWOT/SMOS, and Air-MOSS, as well as global large-sample hydrology in situ data such as Caravan (Kratzert et al., 2023), the GCN250 runoff application (Sujud and Jaafar, 2022), and Global Runoff Data Centre (GRDC) runoff data, focusing on the robustness of groundwater parameters with the relative importance of each hydrologic variable identified.

In addition to identifying the relationship between data and recession parameters, directly utilizing baseflow/streamflow estimates available at the catchment resolution, which enable the catchment-scale estimation of net subsurface discharge (NSD), can also be considered. For example, the recent launch of the SWOT (Surface Water and Ocean Topography) spaceborne mission offers the potential to gather river discharge and baseflow data at a temporal resolution of interest (e.g., daily), with unprecedented global coverage (Baratelli et al., 2018; Li et al., 2020; Wongchuig-Correa et al., 2020). These studies aimed to overcome the spatial and temporal gaps in SWOT observations by integrating a large-scale hydrologic model with synthetic SWOT data through the assimilation of SWOT altimetry data. Based on existing findings about the utility of SWOT data for baseflow estimation, the possibility of obtaining streamflow/baseflow at the catchment resolution with a high temporal resolution (e.g., daily) will be further explored. Hong and Mohanty (2023a) showed that the NSD data could be a key hydrologic variable for the inference of the Boussinesq aquifer's diffusivity downstream of the river. The NSD data provide information on the net mass balance during drought flow by isolating baseflow component using catchments' inflow and outflow data. To this end, we will explore the data-aided streamflow estimates at the catchment resolution and use them as surrogate data (for actual streamflow measurements) to establish the catchment-scale NSD estimates with large-scale coverage. Since all these methods utilize observationally derived data in parameterizations, the emergent/dynamic properties of groundwater can be effectively represented in the framework of LM4-SHARC.

5 Conclusions

The new framework LM4-SHARC presented here harnesses the parametric efficiency of the DF-approximation-based Boussinesq groundwater in capturing the emergent properties of the catchment-scale groundwater. This study proposed a calibration method for the catchment-scale groundwater

based on the accuracy of baseflow fluxes and also demonstrated the contribution of the additional groundwater domain/processes (with tuned groundwater parameters) to the improvements in the catchment-scale water/energy budgets. The streamflow recession analysis provides a physically explicit and viable way to use readily available streamflow measurements to infer the time-evolving groundwater properties. Thus, we argued how the time-evolving groundwater diffusivity could be considered in Earth system modeling through the combined use of a numerical/explicit groundwater domain and the observationally derived stream discharge information.

The notable improvement in soil moisture and temperature predictions resulting from the LM4-SHARC's hydraulic continuum scheme underscores the necessity of resolving the catchment-scale groundwater dynamics and its interactions with the soil and the stream at the grid scale of the ESM. Specifically, our analysis shows that vertical soil drainage to relatively deep groundwater should be taken into account when simulating soil moisture and temperature. We also note that the soil columns in ESMs might hold too much water without the groundwater-induced drainage dynamics. The significant amounts of facilitated drainage (i.e., ΔSWS), roughly around 90–110 mm yr⁻¹, which corresponds to about 6%–7% of the total precipitation at our study site, also emphasizes the importance of considering the lateral groundwater divergence (from the hilltop) and convergence (to the riparian zone) in ESM land components. The existing biases in soil moisture and surface temperature can lead to a flawed description of other land variables, impacting the surface energy balance, carbon cycle, and biogeochemistry. As the simulated soil temperature was found to be lower than observed values due to the ratio of sensible heat to latent heat and soil heat capacity as a function of SMC, liquid water contained in the partially saturated soil needs to be better quantified, as it significantly influences the dynamics of land surface energy balance.

Scaling the fine-scale surface water–energy processes to the global climate model (GCM) grid cell scales while properly considering the hydrologic interactions and heterogeneity is one of the primary objectives of the ESM community. Considering that the streamflow is a major water flux (that significantly affects energy, carbon, and biogeochemical fluxes) crossing the catchments, in order to properly scale the effects of SHARC's catchment-scale hydraulic continuum to the macroscale grid cell, a reach-to-reach connection throughout the river network (i.e., stream/river routing) needs to be established. Also, based on the enhanced baseflow production in LM4-SHARC, we expect significant qualitative enhancements of streamflow estimates, which will, in turn, lead to enhanced surface/near-surface water and energy budgets as well as flooding representation (e.g., floodplain dynamics). Overall, the improved water and energy balance in LM4-SHARC is expected to be relevant for coupled land–atmosphere simulations, where refining the land surface state

plays a significant role in developing the lower-atmospheric boundary layer and also contributes to the efforts to address societal challenges produced by extreme hydrologic events, such as flooding, using enhanced streamflow production.

Appendix A: The dynamic heat capacity of the groundwater and vertical heat conduction (δv_{cnd}) and advection (δv_{adv}) between the soil column and groundwater

LM4-SHARC simulates the heat capacity of the groundwater (hc_{gw}) dynamically according to liquid (i.e., soil moisture) (wl_{gw}) and frozen water content (i.e., soil ice) (ws_{gw}) in the groundwater. LM4-SHARC represents two mechanisms of heat transfer, advection (δv_{adv}) and conduction (δv_{cnd}), between the soil column and the groundwater based on the dynamically estimated groundwater heat capacity.

$$\delta v_{\text{cnd},i}^j = \lambda_{\text{stog},i}^j (T_{\text{gw},i}^j - T_{\text{btm},i}^j) \quad (\text{A1})$$

$$\delta v_{\text{adv},i}^j = \begin{cases} \frac{r_i^j}{\rho} (T_{\text{btm},i}^j - T_{\text{freeze}}) \frac{hc_{\text{gw},i}^j}{\Delta t \Delta L_i^j} & (r > 0) \\ \frac{r_i^j}{\rho} (T_{\text{gw},i}^j - T_{\text{freeze}}) \frac{hc_{\text{gw},i}^j}{\Delta t \Delta L_i^j} & (r \leq 0), \end{cases} \quad (\text{A2})$$

where $\delta v_{\text{cnd},i}^j$ is the vertical heat conduction flux, and $\delta v_{\text{adv},i}^j$ is the advected heat flux between the soil column and groundwater ($\text{J m}^{-2} \text{s}^{-1}$). $\lambda_{\text{stog},i}^j$ is the thermal transmittance ($\text{W m}^{-2} \text{K}^{-1}$) between the bottommost soil layer and groundwater at the i th height band at the j th time step. $T_{\text{gw},i}^j - T_{\text{btm},i}^j$ is the temperature difference between the bottommost layer and the groundwater.

$$hc_{\text{gw},i}^j = hc_{\text{btm},i}^j \times (e_{\text{btm},i} - hl_i \tan \theta) + clw \times wl_{\text{gw},i}^j + csw \times ws_{\text{gw},i}^j, \quad (\text{A3})$$

where wl_{gw} and ws_{gw} are the liquid water and ice content of the groundwater, respectively (–). $hc_{\text{btm},i}^j$ is the dry-soil heat capacity of the bottommost soil layer at the i th height band at the j th time step ($\text{J K}^{-1} \text{m}^{-2}$). clw is the specific heat of liquid water ($4218.0 \text{ J kg}^{-1} \text{K}^{-1}$), and csw is the specific heat of ice ($2106.0 \text{ J kg}^{-1} \text{K}^{-1}$). We note that each HB-scale groundwater domain's thermal properties were assumed to be identical to those of the corresponding soil column's bottommost soil layer.

Code and data availability. The source code for LM4-SHARC v1.0 and the model input data, such as the model domain dataset and forcing data, are shared in a public repository: <https://doi.org/10.5281/zenodo.13750071> (Hong et al., 2024).

Supplement. The supplement related to this article is available online at <https://doi.org/10.5194/gmd-18-2275-2025-supplement>.

Author contributions. All authors contributed to research design and paper editing. MH developed the software, performed the model evaluation and data analysis, and drafted the first version of this paper. NC, SM, and ES contributed to software development. EZ and AP contributed to data preparation.

Competing interests. At least one of the (co-)authors is a member of the editorial board of *Geoscientific Model Development*. The peer-review process was guided by an independent editor, and the authors also have no other competing interests to declare.

Disclaimer. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration, or the US Department of Commerce.

Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. This report was prepared by Minki Hong under award no. NA23OAR4050431I from the National Oceanic and Atmospheric Administration, US Department of Commerce. The authors thank John Dunne and Kirsten Findell at NOAA GFDL for reviewing a first draft of the paper.

Financial support. This research has been supported by the National Oceanic and Atmospheric Administration (grant no. NA23OAR4050431I).

Review statement. This paper was edited by Jeffrey Neal and reviewed by two anonymous referees.

References

- Andreae, M. O., Artaxo, P., Brandao, C., Carswell, F. E., Ciccioli, P., Da Costa, A., Culf, A. D., Esteves, J., Gash, J. H., and Grace, J.: Biogeochemical cycling of carbon, water, energy, trace gases, and aerosols in Amazonia: The LBA-EUSTACH experiments, *J. Geophys. Res.-Atmos.*, 107, LBA33-31–LBA33-25, 2002.
- Bales, R., Stacy, E., Safeeq, M., Meng, X., Meadows, M., Oroza, C., Conklin, M., Glaser, S., and Wagenbrenner, J.: Spatially distributed water-balance and meteorological data from the rain-snow transition, southern Sierra Nevada, California, *Earth Syst. Sci. Data*, 10, 1795–1805, <https://doi.org/10.5194/essd-10-1795-2018>, 2018.
- Baratelli, F., Flipo, N., Rivière, A., and Biancamaria, S.: Retrieving river baseflow from SWOT spaceborne mission, *Remote Sens. Environ.*, 218, 44–54, 2018.
- Baroni, G., Schalge, B., Rakovec, O., Kumar, R., Schüler, L., Samaniego, L., Simmer, C., and Attinger, S.: A Comprehensive Distributed Hydrological Modeling Intercomparison to Support Process Representation and Data Collection Strategies, *Water Resour. Res.*, 55, 990–1010, <https://doi.org/10.1029/2018wr023941>, 2019.
- Bart, R. and Hope, A.: Inter-seasonal variability in baseflow recession rates: The role of aquifer antecedent storage in central California watersheds, *J. Hydrol.*, 519, 205–213, 2014.
- Basha, H. A.: Traveling wave solution of the Boussinesq equation for groundwater flow in horizontal aquifers, *Water Resour. Res.*, 49, 1668–1679, <https://doi.org/10.1002/wrcr.20168>, 2013.
- Best, M. J., Pryor, M., Clark, D., Rooney, G. G., Essery, R., Ménard, C., Edwards, J., Hendry, M., Porson, A., and Gedney, N.: The Joint UK Land Environment Simulator (JULES), model description – Part 1: energy and water fluxes, *Geosci. Model Dev.*, 4, 677–699, 2011.
- Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant, *Hydrolog. Sci. J.*, 24, 43–69, 1979.
- Bisht, G., Huang, M., Zhou, T., Chen, X., Dai, H., Hammond, G. E., Riley, W. J., Downs, J. L., Liu, Y., and Zachara, J. M.: Coupling a three-dimensional subsurface flow and transport model with a land surface model to simulate stream-aquifer-land interactions (CP v1.0), *Geosci. Model Dev.*, 10, 4539–4562, <https://doi.org/10.5194/gmd-10-4539-2017>, 2017.
- Blyth, E. M., Arora, V. K., Clark, D. B., Dadson, S. J., De Kauwe, M. G., Lawrence, D. M., Melton, J. R., Pongratz, J., Turton, R. H., Yoshimura, K., and Yuan, H.: Advances in Land Surface Modelling, *Curr. Clim. Change Rep.*, 7, 45–71, <https://doi.org/10.1007/s40641-021-00171-5>, 2021.
- Boussinesq, J.: Sur le de' bit, en temps de se'cheresse, d'une source alimente' e par une nappe d'eaux d'infiltration, *CR Hebd Acad. Sci.*, 136, 1511–1517, 1903.
- Brutsaert, W.: The unit response of groundwater outflow from a hillslope, *Water Resour. Res.*, 30, 2759–2763, <https://doi.org/10.1029/94wr01396>, 1994.
- Brutsaert, W. and Lopez, J. P.: Basin-scale geohydrologic drought flow features of riparian aquifers in the Southern Great Plains, *Water Resour. Res.*, 34, 233–240, <https://doi.org/10.1029/97wr03068>, 1998.
- Brutsaert, W. and Nieber, J. L.: Regionalized drought flow hydrographs from a mature glaciated plateau, *Water Resour. Res.*, 13, 637–643, <https://doi.org/10.1029/wr013i003p00637>, 1977.
- Cai, H., Shi, H., Liu, S., and Babovic, V.: Impacts of regional characteristics on improving the accuracy of groundwater level prediction using machine learning: The case of central eastern continental United States, *J. Hydrol.: Reg. Stud.*, 37, 100930, <https://doi.org/10.1016/j.ejrh.2021.100930>, 2021.
- Campoy, A., Ducharme, A., Cheruy, F., Hourdin, F., Polcher, J., and Dupont, J.: Response of land surface fluxes and precipitation to different soil bottom hydrological conditions in a general circulation model, *J. Geophys. Res.-Atmos.*, 118, 10725–10739, 2013.
- Chaney, N. W., Metcalfe, P., and Wood, E. F.: HydroBlocks: a field-scale resolving land surface model for application over continental extents, *Hydrol. Process.*, 30, 3543–3559, <https://doi.org/10.1002/hyp.10891>, 2016.

- Chaney, N. W., Van Huijgevoort, M. H. J., Shevliakova, E., Malyshev, S., Milly, P. C. D., Gauthier, P. P. G., and Sulman, B. N.: Harnessing big data to rethink land heterogeneity in Earth system models, *Hydrol. Earth Syst. Sci.*, 22, 3311–3330, <https://doi.org/10.5194/hess-22-3311-2018>, 2018.
- Chaney, N. W., Torres-Rojas, L., Vergopolan, N., and Fisher, C. K.: HydroBlocks v0.2: enabling a field-scale two-way coupling between the land surface and river networks in Earth system models, *Geosci. Model Dev.*, 14, 6813–6832, <https://doi.org/10.5194/gmd-14-6813-2021>, 2021.
- Chen, X. and Hu, Q.: Groundwater influences on soil moisture and surface evaporation, *J. Hydrol.*, 297, 285–300, 2004.
- Cheng, S., Cheng, L., Liu, P., Qin, S., Zhang, L., Xu, C. Y., Xiong, L., Liu, L., and Xia, J.: An Analytical Baseflow Coefficient Curve for Depicting the Spatial Variability of Mean Annual Catchment Baseflow, *Water Resour. Res.*, 57, e2020WR029529, <https://doi.org/10.1029/2020wr029529>, 2021.
- Clark, M. P., Rupp, D. E., Woods, R. A., Tromp-Van Meerveld, H. J., Peters, N. E., and Freer, J. E.: Consistency between hydrological models and field observations: linking processes at the hillslope scale to hydrological responses at the watershed scale, *Hydrol. Process.*, 23, 311–319, <https://doi.org/10.1002/hyp.7154>, 2009.
- Clark, M. P., Fan, Y., Lawrence, D. M., Adam, J. C., Bolster, D., Gochis, D. J., Hooper, R. P., Kumar, M., Leung, L. R., Mackay, D. S., Maxwell, R. M., Shen, C. P., Swenson, S. C., and Zeng, X. B.: Improving the representation of hydrologic processes in Earth System Models, *Water Resour. Res.*, 51, 5929–5956, <https://doi.org/10.1002/2015wr017096>, 2015.
- De Rosnay, P., Polcher, J. D., Bruen, M., and Laval, K.: Impact of a physically based soil water flow and soil-plant interaction representation for modeling large-scale land surface processes, *J. Geophys. Res.-Atmos.*, 107, ACL3-1–ACL3-19, 2002.
- De St Venant, B.: Theorie du mouvement non-permanent des eaux avec application aux crues des rivières et à l'introduction des Mares dans leur lit, *Academic de Sci. Comptes Redus*, 73, 148–154, 1871.
- Delworth, T. L., Cooke, W. F., Adcroft, A., Bushuk, M., Chen, J. H., Dunne, K. A., Ginoux, P., Gudgel, R., Hallberg, R. W., and Harris, L.: SPEAR: The next generation GFDL modeling system for seasonal to multidecadal prediction and projection, *J. Adv. Model. Earth Sy.*, 12, e2019MS001895, <https://doi.org/10.1029/2019MS001895>, 2020.
- Dunne, J. P., Horowitz, L., Adcroft, A., Ginoux, P., Held, I., John, J., Krasting, J. P., Malyshev, S., Naik, V., and Paulot, F.: The GFDL Earth System Model version 4.1 (GFDL-ESM 4.1): Overall coupled model description and simulation characteristics, *J. Adv. Model. Earth Sy.*, 12, e2019MS002015, <https://doi.org/10.1029/2019MS002015>, 2020.
- Dupuit, J.: Etudes theoriques et pratiques sur le mouvement des eaux dans les canaux decouverts et a travers les terrains permeables avec des considerations relatives au regime des grandes eaux, au debouche a leur donner et a la marche des alluvions dans les rivieres a fond mobile, Dunod, https://archive.org/details/bub_gb_NDK90LcYwhEC/page/n149/mode/2up (last access: 14 March 2024), 1863.
- Fan, Y.: Groundwater in the Earth's critical zone: Relevance to large-scale patterns and processes, *Water Resour. Res.*, 51, 3052–3069, <https://doi.org/10.1002/2015wr017037>, 2015.
- Fan, Y. and Miguez-Macho, G.: A simple hydrologic framework for simulating wetlands in climate and earth system models, *Clim. Dynam.*, 37, 253–278, 2011.
- Fan, Y., Miguez-Macho, G., Weaver, C. P., Walko, R., and Robock, A.: Incorporating water table dynamics in climate modeling: 1. Water table observations and equilibrium water table simulations, *J. Geophys. Res.-Atmos.*, 112, D10125, <https://doi.org/10.1029/2006jd008111>, 2007.
- Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., Brooks, P. D., Dietrich, W. E., Flores, A., Grant, G., Kirchner, J. W., Mackay, D. S., McDonnell, J. J., Milly, P. C. D., Sullivan, P. L., Tague, C., Ajami, H., Chaney, N., Hartmann, A., Hazenberg, P., McNamara, J., Pelletier, J., Perket, J., Rouholahnejad-Freund, E., Wagener, T., Zeng, X., Beighley, E., Buzan, J., Huang, M., Livneh, B., Mohanty, B. P., Nijssen, B., Safeeq, M., Shen, C., Verseveld, W., Volk, J., and Yamazaki, D.: Hillslope Hydrology in Global Change Research and Earth System Modeling, *Water Resour. Res.*, 55, 1737–1772, <https://doi.org/10.1029/2018wr023903>, 2019.
- Forchheimer, P.: Über die ergiebigkeit von brummen, Anlagen und Sickerschlitzten, *Zeitsch. Archit. Ing. Ver.*, Hannover, 32, 539–563, 1986.
- Gentine, P., Green, J. K., Guérin, M., Humphrey, V., Seneviratne, S. I., Zhang, Y., and Zhou, S.: Coupling between the terrestrial carbon and water cycles – a review, *Environ. Res. Lett.*, 14, 083003, <https://doi.org/10.1088/1748-9326/ab22d6>, 2019.
- Gleeson, T., Wang-Erlandsson, L., Porkka, M., Zipper, S. C., Jaramillo, F., Gerten, D., Fetzer, I., Cornell, S. E., Piemontese, L., Gordon, L. J., Rockström, J., Oki, T., Sivapalan, M., Wada, Y., Brauman, K. A., Flörke, M., Bierkens, M. F. P., Lehner, B., Keys, P., Kummu, M., Wagener, T., Dadson, S., Troy, T. J., Steffen, W., Falkenmark, M., and Famiglietti, J. S.: Illuminating water cycle modifications and Earth system resilience in the Anthropocene, *Water Resour. Res.*, 56, e2019WR024957, <https://doi.org/10.1029/2019wr024957>, 2020.
- Gochis, D. J., Barlage, M., Dugger, A., FitzGerald, K., Karsten, L., McAllister, M., McCreight, J., Mills, J., RafieeiNasab, A., Read, L., Sampson, K., Yates, D., and Yu, W.: The NCAR WRF-Hydro Modeling System Technical Description (Version 5.2), National Center for Atmospheric Research, <https://ral.ucar.edu/sites/default/files/docs/water/wrf-hydro-v511-technical-description.pdf> (last access: 1 May 2024), 2021.
- Gómez-Hernández, J. J. and Gorelick, S. M.: Effective groundwater model parameter values: Influence of spatial variability of hydraulic conductivity, leakance, and recharge, *Water Resour. Res.*, 25, 405–419, 1989.
- Heath, R. C.: Basic ground-water hydrology, US Geological Survey, <https://pubs.usgs.gov/wsp/2220/report.pdf> (last access: 22 April 2024), 2004.
- Hong, M. and Mohanty, B.: A new method for effective parameterization of catchment-scale aquifer through event-scale recession analysis, *Adv. Water Resour.*, 174, 104408, <https://doi.org/10.1016/j.advwatres.2023.104408>, 2023a.
- Hong, M. and Mohanty, B. P.: Representing Bidirectional Hydraulic Continuum Between the Stream and Hillslope in the National Water Model for Improved Streamflow Prediction, *J. Adv. Model. Earth Sy.*, 15, e2022MS003325, <https://doi.org/10.1029/2022ms003325>, 2023b.

- Hong, M., Chaney, N., Malyshev, S., and Shevliakova, E.: LM4-SHARC v1.0: Resolving the Catchment-scale Soil-Hillslope Aquifer-River Continuum for the GFDL Earth System Modeling Framework, Zenodo [code and data set], <https://doi.org/10.5281/zenodo.13750071>, 2024.
- Hornberger, G. M. and Remson, I.: A moving boundary model of a one-dimensional saturated-unsaturated, transient porous flow system, *Water Resour. Res.*, 6, 898–905, <https://doi.org/10.1029/WR006i003p00898>, 1970.
- Hunsaker, C. T., Whitaker, T. W., and Bales, R. C.: Snowmelt Runoff and Water Yield Along Elevation and Temperature Gradients in California's Southern Sierra Nevada, *J. Am. Water Resour. As.*, 48, 667–678, <https://doi.org/10.1111/j.1752-1688.2012.00641.x>, 2012.
- Jachens, E. R., Rupp, D. E., Roques, C., and Selker, J. S.: Recession analysis revisited: impacts of climate on parameter estimation, *Hydrol. Earth Syst. Sci.*, 24, 1159–1170, <https://doi.org/10.5194/hess-24-1159-2020>, 2020.
- Jepsen, S. M., Harmon, T. C., Meadows, M. W., and Hunsaker, C. T.: Hydrogeologic influence on changes in snowmelt runoff with climate warming: Numerical experiments on a mid-elevation catchment in the Sierra Nevada, USA, *J. Hydrol.*, 533, 332–342, 2016.
- Jing, M., Heße, F., Kumar, R., Kolditz, O., Kalbacher, T., and Attinger, S.: Influence of input and parameter uncertainty on the prediction of catchment-scale groundwater travel time distributions, *Hydrol. Earth Syst. Sci.*, 23, 171–190, 2019.
- Karlsen, R. H., Bishop, K., Grabs, T., Ottosson-Löfvenius, M., Laudon, H., and Seibert, J.: The role of landscape properties, storage and evapotranspiration on variability in streamflow recessions in a boreal catchment, *J. Hydrol.*, 570, 315–328, <https://doi.org/10.1016/j.jhydrol.2018.12.065>, 2019.
- Kirchner, J. W.: Catchments as simple dynamical systems: Catchment characterization, rainfall–runoff modeling, and doing hydrology backward, *Water Resour. Res.*, 45, W02429, <https://doi.org/10.1029/2008wr006912>, 2009.
- Kollet, S. J. and Maxwell, R. M.: Integrated surface-groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model, *Adv. Water Resour.*, 29, 945–958, <https://doi.org/10.1016/j.advwatres.2005.08.006>, 2006.
- Kollet, S. J. and Maxwell, R. M.: Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model, *Water Resour. Res.*, 44, <https://doi.org/10.1029/2007WR006004>, 2008.
- Kratzert, F., Nearing, G., Addor, N., Erickson, T., Gauch, M., Gilon, O., Gudmundsson, L., Hassidim, A., Klotz, D., Nevo, S., Shalev, G., and Matias, Y.: Caravan – A global community dataset for large-sample hydrology, *Sci. Data*, 10, 61, <https://doi.org/10.1038/s41597-023-01975-w>, 2023.
- Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang, Z. L., Levis, S., and Sakaguchi, K.: Parameterization improvements and functional and structural advances in version 4 of the Community Land Model, *J. Adv. Model. Earth Sy.*, 3, M03001, <https://doi.org/10.1029/2011MS00045>, 2011.
- Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., Broeke, M., Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Val Martin, M., and Zeng, X.: The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty, *J. Adv. Model. Earth Sy.*, 11, 4245–4287, <https://doi.org/10.1029/2018ms001583>, 2019.
- Leung, L. R., Huang, M., Qian, Y., and Liang, X.: Climate–soil–vegetation control on groundwater table dynamics and its feedbacks in a climate model, *Clim. Dynam.*, 36, 57–81, 2011.
- Li, D., Andreadis, K. M., Margulis, S. A., and Lettenmaier, D. P.: A Data Assimilation Framework for Generating Space–Time Continuous Daily SWOT River Discharge Data Products, *Water Resour. Res.*, 56, e2019WR026999, <https://doi.org/10.1029/2019wr026999>, 2020.
- Li, H. and Ameli, A.: A statistical approach for identifying factors governing streamflow recession behaviour, *Hydrol. Process.*, 36, e14718, <https://doi.org/10.1002/hyp.14718>, 2022.
- Li, H., Wigmosta, M. S., Wu, H., Huang, M., Ke, Y., Coleman, A. M., and Leung, L. R.: A physically based runoff routing model for land surface and earth system models, *J. Hydrometeorol.*, 14, 808–828, 2013.
- Liang, X., Xie, Z. H., and Huang, M. Y.: A new parameterization for surface and groundwater interactions and its impact on water budgets with the variable infiltration capacity (VIC) land surface model, *J. Geophys. Res.–Atmos.*, 108, 8613, <https://doi.org/10.1029/2002jd003090>, 2003.
- Manning, R.: On the flow of water in open channels and pipes, *Transactions of the Institution of Civil Engineers of Ireland*, 20, 161–207, 1891.
- Maxwell, R. M. and Kollet, S. J.: Interdependence of groundwater dynamics and land-energy feedbacks under climate change, *Nat. Geosci.*, 1, 665–669, <https://doi.org/10.1038/ngeo315>, 2008.
- Maxwell, R. M., Chow, F. K., and Kollet, S. J.: The groundwater–land-surface–atmosphere connection: Soil moisture effects on the atmospheric boundary layer in fully-coupled simulations, *Adv. Water Resour.*, 30, 2447–2466, <https://doi.org/10.1016/j.advwatres.2007.05.018>, 2007.
- Maxwell, R. M., Lundquist, J. K., Mirocha, J. D., Smith, S. G., Woodward, C. S., and Tompson, A. F. B.: Development of a Coupled Groundwater–Atmosphere Model, *Mon. Weather Rev.*, 139, 96–116, <https://doi.org/10.1175/2010mwr3392.1>, 2011.
- Maxwell, R. M., Condon, L. E., and Kollet, S. J.: A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3, *Geosci. Model Dev.*, 8, 923–937, <https://doi.org/10.5194/gmd-8-923-2015>, 2015.
- McCabe, M., Wood, E. F., Wójcik, R., Pan, M., Sheffield, J., Gao, H., and Su, H.: Hydrological consistency using multi-sensor remote sensing data for water and energy cycle studies, *Remote Sens. Environ.*, 112, 430–444, 2008.
- Miguez-Macho, G. and Fan, Y.: The role of groundwater in the Amazon water cycle: 1. Influence on seasonal streamflow,

- flooding and wetlands, *J. Geophys. Res.-Atmos.*, 117, D15113, <https://doi.org/10.1029/2012jd017539>, 2012.
- Miguez-Macho, G., Fan, Y., Weaver, C. P., Walko, R., and Robock, A.: Incorporating water table dynamics in climate modeling: 2. Formulation, validation, and soil moisture simulation, *J. Geophys. Res.-Atmos.*, 112, D13108, <https://doi.org/10.1029/2006jd008112>, 2007.
- Milly, P. C., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., and Stouffer, R. J.: Stationarity is dead: Whither water management?, *Science*, 319, 573–574, 2008.
- Milly, P. C. D., Malyshev, S. L., Shevliakova, E., Dunne, K. A., Findell, K. L., Gleeson, T., Liang, Z., Philipps, P., Stouffer, R. J., and Swenson, S.: An Enhanced Model of Land Water and Energy for Global Hydrologic and Earth-System Studies, *J. Hydrometeorol.*, 15, 1739–1761, <https://doi.org/10.1175/jhm-d-13-0162.1>, 2014.
- Mu, Q., Zhao, M., and Running, S. W.: Evolution of hydrological and carbon cycles under a changing climate, *Hydrol. Process.*, 25, 4093–4102, <https://doi.org/10.1002/hyp.8367>, 2011.
- Newman, A. J., Clark, M. P., Winstral, A., Marks, D., and Seyfried, M.: The Use of Similarity Concepts to Represent Subgrid Variability in Land Surface Models: Case Study in a Snowmelt-Dominated Watershed, *J. Hydrometeorol.*, 15, 1717–1738, <https://doi.org/10.1175/jhm-d-13-038.1>, 2014.
- Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., and Xia, Y.: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, *J. Geophys. Res.*, 116, D12109, <https://doi.org/10.1029/2010jd015139>, 2011.
- Oleson, K. W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., Levis, S., Li, F., Riley, W. J., Subin, Z. M., Swenson, S. C., and Thornton, P. E.: Technical Description of version 4.5 of the Community Land Model (CLM), NCAR Earth System Laboratory, <https://doi.org/10.5065/D6RR1W7M>, 2013.
- Ouyang, L., Liu, S., Ye, J., Liu, Z., Sheng, F., Wang, R., and Lu, Z.: Quantitative assessment of surface runoff and base flow response to multiple factors in Pengchongjian small watershed, *Forests*, 9, 553, <https://doi.org/10.3390/f9090553>, 2018.
- Pappenberger, F., Dutra, E., Wetterhall, F., and Cloke, H. L.: Deriving global flood hazard maps of fluvial floods through a physical model cascade, *Hydrol. Earth Syst. Sc.*, 16, 4143–4156, 2012.
- Pauritsch, M., Birk, S., Wagner, T., Hergarten, S., and Winkler, G.: Analytical approximations of discharge recessions for steeply sloping aquifers in alpine catchments, *Water Resour. Res.*, 51, 8729–8740, <https://doi.org/10.1002/2015wr017749>, 2015.
- Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., and Rossiter, D.: SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty, *SOIL*, 7, 217–240, <https://doi.org/10.5194/soil-7-217-2021>, 2021.
- Pokhrel, Y. N., Fan, Y., Miguez-Macho, G., Yeh, P. J. F., and Han, S.-C.: The role of groundwater in the Amazon water cycle: 3. Influence on terrestrial water storage computations and comparison with GRACE, *J. Geophys. Res.-Atmos.*, 118, 3233–3244, <https://doi.org/10.1002/jgrd.50335>, 2013.
- Rupp, D. E. and Selker, J. S.: On the use of the Boussinesq equation for interpreting recession hydrographs from sloping aquifers, *Water Resour. Res.*, 42, W12421, <https://doi.org/10.1029/2006wr005080>, 2006.
- Safeeq, M. and Hunsaker, C. T.: Characterizing Runoff and Water Yield for Headwater Catchments in the Southern Sierra Nevada, *J. Am. Water Resour. As.*, 52, 1327–1346, <https://doi.org/10.1111/1752-1688.12457>, 2016.
- Shangguan, W., Dai, Y., Duan, Q., Liu, B., and Yuan, H.: A global soil data set for earth system modeling, *J. Adv. Model. Earth Sy.*, 6, 249–263, <https://doi.org/10.1002/2013ms000293>, 2014.
- Shen, C. P. and Phanikumar, M. S.: A process-based, distributed hydrologic model based on a large-scale method for surface-subsurface coupling, *Adv. Water Resour.*, 33, 1524–1541, <https://doi.org/10.1016/j.advwatres.2010.09.002>, 2010.
- Shevliakova, E., Malyshev, S., Martinez-Cano, I., Milly, P. C. D., Pacala, S. W., Ginoux, P., Dunne, K. A., Dunne, J. P., Dupuis, C., Findell, K. L., Ghannam, K., Horowitz, L. W., Knutson, T. R., Krasting, J. P., Naik, V., Philipps, P., Zadeh, N., Yu, Y., Zeng, F., and Zeng, Y.: The Land Component LM4.1 of the GFDL Earth System Model ESM4.1: Model Description and Characteristics of Land Surface Climate and Carbon Cycling in the Historical Simulation, *J. Adv. Model. Earth Sy.*, 16, e2023MS003922, <https://doi.org/10.1029/2023ms003922>, 2024.
- Sivapalan, M.: Pattern, process and function: elements of a unified theory of hydrology at the catchment scale, *Encyclopedia of hydrological sciences, Encyclopedia of Hydrological Sciences*, <https://doi.org/10.1002/0470848944.hsa012>, 2006.
- Strack, O. D. L.: A Dupuit-Forchheimer Model for three-dimensional flow with variable density, *Water Resour. Res.*, 31, 3007–3017, <https://doi.org/10.1029/95wr02254>, 1995.
- Strelkoff, T.: Numerical solution of Saint-Venant equations, *J. Hydraul. Div.*, 96, 223–252, 1970.
- Subin, Z. M., Milly, P. C. D., Sulman, B. N., Malyshev, S., and Shevliakova, E.: Resolving terrestrial ecosystem processes along a subgrid topographic gradient for an earth-system model, *Hydrol. Earth Syst. Sci. Discuss.*, 11, 8443–8492, <https://doi.org/10.5194/hessd-11-8443-2014>, 2014.
- Sujud, L. H. and Jaafar, H. H.: A global dynamic runoff application and dataset based on the assimilation of GPM, SMAP, and GCN250 curve number datasets, *Sci. Data*, 9, 706, <https://doi.org/10.1038/s41597-022-01834-0>, 2022.
- Szilagyi, J. and Parlange, M. B.: Baseflow separation based on analytical solutions of the Boussinesq equation, *J. Hydrol.*, 204, 251–260, [https://doi.org/10.1016/s0022-1694\(97\)00132-7](https://doi.org/10.1016/s0022-1694(97)00132-7), 1998.
- Szilagyi, J., Parlange, M. B., and Albertson, J. D.: Recession flow analysis for aquifer parameter determination, *Water Resour. Res.*, 34, 1851–1857, <https://doi.org/10.1029/98wr01009>, 1998.
- Takata, K., Emori, S., and Watanabe, T.: Development of the minimal advanced treatments of surface interaction and runoff, *Global Planet. Change*, 38, 209–222, 2003.
- Tashie, A., Pavelsky, T., and Band, L. E.: An Empirical Reevaluation of Streamflow Recession Analysis at the Continental Scale, *Water Resour. Res.*, 56, e2019WR025448, <https://doi.org/10.1029/2019wr025448>, 2020.
- Tashie, A., Pavelsky, T., Band, L., and Topp, S.: Watershed-Scale Effective Hydraulic Properties of the Continental United

- States, J. *Adv. Model. Earth Sy.*, 13, e2020MS002440, <https://doi.org/10.1029/2020ms002440>, 2021.
- Troch, P. A., De Troch, F. P., and Brutsaert, W.: Effective water table depth to describe initial conditions prior to storm rainfall in humid regions, *Water Resour. Res.*, 29, 427–434, <https://doi.org/10.1029/92wr02087>, 1993.
- Troch, P. A., Berne, A., Bogaart, P., Harman, C., Hilberts, A. G. J., Lyon, S. W., Paniconi, C., Pauwels, V. R. N., Rupp, D. E., Selker, J. S., Teuling, A. J., Uijlenhoet, R., and Verhoest, N. E. C.: The importance of hydraulic groundwater theory in catchment hydrology: The legacy of Wilfried Brutsaert and Jean-Yves Parlange, *Water Resour. Res.*, 49, 5099–5116, <https://doi.org/10.1002/wrcr.20407>, 2013.
- Trotter, L., Saft, M., Peel, M. C., and Fowler, K. J.: Recession constants are non-stationary: Impacts of multi-annual drought on catchment recession behaviour and storage dynamics, *J. Hydrol.*, 630, 130707, <https://doi.org/10.1016/j.jhydrol.2024.130707>, 2024.
- USGS: 3D Elevation Program 1-Meter Resolution Digital Elevation Model [data set], <https://apps.nationalmap.gov/downloader/> (last access: 1 September 2023), 2019.
- Vannier, O., Braud, I., and Anquetin, S.: Regional estimation of catchment-scale soil properties by means of streamflow recession analysis for use in distributed hydrological models, *Hydrol. Process.*, 28, 6276–6291, <https://doi.org/10.1002/hyp.10101>, 2014.
- Wongchuig-Correa, S., de Paiva, R. C. D., Biancamaria, S., and Collischonn, W.: Assimilation of future SWOT-based river elevations, surface extent observations and discharge estimations into uncertain global hydrological models, *J. Hydrol.*, 590, 125473, <https://doi.org/10.1016/j.jhydrol.2020.125473>, 2020.
- Xu, B., Ye, M., Dong, S., Dai, Z., and Pei, Y.: A new model for simulating spring discharge recession and estimating effective porosity of karst aquifers, *J. Hydrol.*, 562, 609–622, 2018.
- Xu, Z., Siirila-Woodburn, E. R., Rhoades, A. M., and Feldman, D.: Sensitivities of subgrid-scale physics schemes, meteorological forcing, and topographic radiation in atmosphere-through-bedrock integrated process models: a case study in the Upper Colorado River basin, *Hydrol. Earth Syst. Sci.*, 27, 1771–1789, 2023.
- Yeh, P. J. F. and Eltahir, E. A. B.: Representation of Water Table Dynamics in a Land Surface Scheme. Part I: Model Development, *J. Climate*, 18, 1861–1880, <https://doi.org/10.1175/jcli3330.1>, 2005.
- Zecharias, Y. B. and Brutsaert, W.: Recession characteristics of groundwater outflow and base flow from mountainous watersheds, *Water Resour. Res.*, 24, 1651–1658, 1988.
- Zeng, Y., Xie, Z., Yu, Y., Liu, S., Wang, L., Jia, B., Qin, P., and Chen, Y.: Ecohydrological effects of stream–aquifer water interaction: a case study of the Heihe River basin, northwestern China, *Hydrol. Earth Syst. Sci.*, 20, 2333–2352, <https://doi.org/10.5194/hess-20-2333-2016>, 2016.
- Zhao, M., Golaz, J. C., Held, I., Guo, H., Balaji, V., Benson, R., Chen, J. H., Chen, X., Donner, L., and Dunne, J.: The GFDL global atmosphere and land model AM4, 0/LM4, 0:2, Model description, sensitivity studies, and tuning strategies, *J. Adv. Model. Earth Sy.*, 10, 735–769, 2018.
- Zorzetto, E., Malyshev, S., Chaney, N., Paynter, D., Menzel, R., and Shevliakova, E.: Effects of complex terrain on the shortwave radiative balance: a sub-grid-scale parameterization for the GFDL Earth System Model version 4.1, *Geosci. Model Dev.*, 16, 1937–1960, <https://doi.org/10.5194/gmd-16-1937-2023>, 2023.