LM4-SHARC v1.0: Resolving the Catchment-scale Soil-Hillslope Aquifer-River Continuum for the GFDL Earth System Modeling Framework

Minki Hong[1], Nathaniel Chaney[2], Sergey Malyshev[3], Enrico Zorzetto[4], Anthony Preucil[3], and Elena Shevliakova[3]

[1] Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA
[2] Department of Civil and Environmental Engineering, Duke University, Durham, NC, USA
[3] NOAA OAR Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA
[4] Department of Earth and Environmental Science, New Mexico Institute of Mining and Technology, NM, USA

Correspondence to: Minki Hong (mh0093@princeton.edu)

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 its implication to the broader Earth system. To fill this gap, we developed a new parameterization scheme for resolving 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 Soil-Hillslope Aquifer-River Continuum), in the Geophysical Fluid Dynamics Laboratory land model (i.e., LM4-SHARC) and evaluated its performance. By bridging the gap between hydraulic groundwater theory and ESMs' land hydrology, the new LM4-SHARC provides a path forward to learn the groundwater's emergent properties from available streamflow data (i.e., recession analysis), enhancing the representation of sub-grid variability of 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 representing 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 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, the 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 is 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 studies commonly found higher sensitivity of surface energy balance to groundwater storage if the water table is shallow (e.g., the water table depth less than 5 m).

However, many Earth system models (ESMs) only represent a few top meters of soil, and most, if not all, ignore two-way interaction between the groundwater and other components of the hydrologic cycle. Specifically, ESMs lack proper land surface scheme(s) to represent the two-way interactions between the groundwater and stream/river while conserving the hydraulic continuity between them. In most cases, the river routing module implemented in the ESMs exists mainly for their 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 groundwater (Li et al., 2013; De Rosnay et al., 2002; Lawrence et al., 2011; Lawrence et al., 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 an atmosphere-land water information produced at the resolution of 0.25° (Delworth et al., 2020) — considered the most high-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 to directly use for locating hydrologic events at a fine-scale even in 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 ESMs' 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 sub-grid scale (e.g., catchment) to adequately represent the sub-grid variability of hydrologic states and its implication 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 sub-grid 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 have been 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, many efforts have 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). These efforts mainly aimed at representing the water-energy coupled balance while considering the time-dependent non-linear relationship between water states and fluxes based on Darcian flow (Clark et al., 2015; Fan et al., 2019).

However, more specifically concerning 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, an 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 movements significantly affect the catchments' water and energy dynamics (Fan, 2015; Miguez-Macho and Fan, 2012; Chen and Hu, 2004; Maxwell et al., 2007). Even in existing studies to capture the interactions among the soil-groundwater-stream, the properties of the groundwater were treated as constant and not as emergent and dynamic properties as the result of transient groundwater storage affected by 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 for a storage-discharge function (e.g., bucket model) or 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 been largely absent in those efforts.

The Dupuit-Forchheimer (DF) approximation allows the incorporation of the catchment-scale groundwater into ESMs and accounts for its emergent properties/processes shaping 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 combined effects of the groundwater properties 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 not representable below model-specific resolution can be considered one of the challenges to 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). The 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 ESMs' grid-scale via approaches to the sub-grid model structure such as HydroBlocks (Chaney et al., 2016; Chaney et al., 2018). The hierarchical multivariate clustering (HMC) approach, as implemented in LM4-HydroBlocks (Chaney et al., 2018), enables partitioning a

macroscale land domain grid (e.g., 0.5°, 1.0°) into sub-grid terrain units (termed 'tiles') (Milly et al., 2014; Zhao et al., 2018; Dunne et al., 2020). This study presents a new parameterization of 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 version 4 (i.e., LM4) aiming to represent the catchment-scale two-way interactions among the unsaturated soil (i.e., vadose zone), the groundwater, and stream/river in the GFDL Earth system modeling framework. The new LM4-SHARC has been developed by extending the existing hillslope hydrology scheme currently used in LM4-HydroBlocks (Subin et al., 2014; Chaney et al., 2018), and relying on the catchment-scale hydrologic partitioning (of the macroscale grid cell) by HMC approach for the explicit treatment of divergence fluxes among soil-groundwater-stream.

In sum, the 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–groundwater and in the lateral direction between the groundwater–stream, and (3) accounts for the groundwater-induced variations in surface water-energy budgets to enhance the ESM's realism of land hydrology. This study primarily compares the two model configurations to evaluate the importance of accurately describing the catchment-scale groundwater in the land component of ESM. As a proof of concept, we apply the LM4-SHARC and LM4-HydroBlocks to a headwater catchment in the Sierra National Forest, Nevada, and evaluate the modeled outputs against corresponding observations. Additionally, we discuss how the hydrologic convergence to the river valley (and divergence from the hilltop) of liquid water and thermal energy due to the groundwater flow 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; Chaney et al., 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 sub-grid 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 previously used in constructing land domains for LM4-HydroBlocks to account for the sub-grid spatial variability of land properties within a regular latitude-longitude climate grid cell (e.g., at 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 hillslope clustered in a set of k characteristic hillslopes (CH). 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 as "height bands" (HB), 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 representative element volume (REV) of soil with respect to covariates incorporated for clustering (Chaney et al., 2018). The channels were delineated using an area threshold of 100,000 m^2. Three parameters k, dh, and p are required in the land pre-processing code; k determines the number of CH in the land model domain, dh defines the height difference

148   between adjacent HBs, thus determining the number of HBs in a CH. The parameter p sets the number of tiles in an

149   HB (i.e., the intra-HB variability). For further details on the HMC algorithm, see Chaney et al (2018) and Chaney et

150   al (2021).

151   Unlike the previous configuration of the land input dataset for a regular grid cell, this study, for the first

152   time, refined the existing HMC framework to generate a land input dataset for a catchment domain with an irregular

153   boundary. This refinement has been designed to assess the GFDL land model's performance at the catchment level

154   while enhancing the interpretability of the land model's hydrologic outputs. The catchment's boundary was

155   determined by computing the geographic extent of the contributing area that drains to a specific point (i.e., the

156   catchment outlet). Since the catchment areas are significantly variable (roughly ranging from sub-km^2 to hundreds

157   of km^2), the resolution of a digital elevation model (DEM) used for clustering the terrain characteristics should be

158   fine enough to capture the intra-catchment heterogeneity of terrain characteristics adequately. To this end, the USGS

159   3DEP 1/3 arc-second (i.e., 10-m resolution) DEM was used in this study (USGS, 2019).

160

161   **2.2. Representation of the soil-groundwater processes in LM4-HydroBlocks**

162   In the current configuration of GFDL LM4-HydroBlocks, the impermeable bedrock is assumed to exist 10

163   m below the ground, and the sub-grid scale (e.g., reach) river dynamics were not represented. The lateral liquid

164   water flux between the saturated soil layers in each adjacent tile was considered groundwater flow. Hence, the

165   groundwater properties depend on surface/near-surface properties from which each soil column's soil properties are

166   derived (Milly et al., 2014; Shevliakova et al., 2024). Each soil column in a catchment is composed of a vertical 1-D

167   model (i.e., soil-bedrock column) from canopy air down to an impermeable bedrock layer (Milly et al., 2014). The

168   processes resolved by the 1-D model include surface energy balance, vegetation dynamics, plant hydraulics,

169   photosynthesis, snow physics, and soil thermal and hydraulic physics (Zorzetto et al., 2023; Subin et al., 2014; Milly

170   et al., 2014). All the soil–bedrock columns are simulated with a 10 m soil depth at a 30-minute physical time step.

171   The soil–bedrock 10-m column is discretized into 20 layers. At the surface, the water flux boundary condition is

172   tentatively set to the difference between the sum of rainfall and snowmelt minus evaporation at each time step (i.e.,

173   time-dependent flow). The heat flux boundary condition at the surface is determined by the balance between the

174   turbulent and the radiative fluxes. At the bottom of the soil–bedrock column (i.e., 10 m below the land surface),

175   water flux is assumed to be zero, implying that the impermeable bedrock is always located 10 m below the land

176   surface. Due to this impermeable layer assumption, the water table is always identified within 10 m of the land

177   surface. A constant geothermal heat flux is prescribed at the bottom of the soil–bedrock columns (Milly et al., 2014).

178   The catchment-scale hydrologic structure within the macroscale grid cell is described based on the inter-tile

179   connection enforced through the flow accumulation area derived from the digital elevation model (DEM) (Chaney et

180   al., 2018). The water fluxes between adjacent tiles (i.e., inter-tile) are simulated on a layer-to-layer basis, defined

181   following the soil layer order from the land surface, according to the gradient of total pressure head between the

182   corresponding soil layers. The inter-tile heat and dissolved organic carbon fluxes advected by water transport are

183   also represented. While the tile located right next to the river reach interacts with the river, the flux from the tile to

184   the river is a one-way flux as the reach-scale streams' (corresponding to a catchment) hydrograph are not accounted

185   for in LM4-HydroBlocks. With respect to the flux exchange between the river and the soil–bedrock column

186   (adjacent to the reach), in fact, the river stage is set to zero (static pressure head in the reach). Consequently, the

187   water flux is invariably one-way, from the hillslope to the river.

188

189   **2.3. Representation of the Soil-Hillslope Aquifer-River Interactions in LM4-SHARC**

190   **2.3.1. Two-way water transport and conservation in LM4-SHARC**

191   The LM4-SHARC solves the non-linear Boussinesq equation, which was derived from the DF

192   approximation to represent the lateral groundwater discharge flux from the hilltop to the river in a homogeneous and

193   isotropic unconfined groundwater (i.e., stream-hillslope interface) (Basha, 2013; Hong and Mohanty, 2023b;

194   Hornberger and Remson, 1970). According to the DF approximation, the unconfined groundwater flows laterally,

195   and the lateral discharge flux is proportional to the saturated groundwater thickness (Dupuit, 1863; Forchheimer,

196   1986). Therefore, the lateral hydraulic gradient is the only driver of groundwater lateral discharge fluxes as the

197   equipotential lines in the saturated zone are assumed to be vertical (i.e., hydrostatic). For saturated groundwater flow

198   in an unconfined groundwater overlaying an impermeable bedrock of slope θ, the lateral groundwater discharge flux

199   is estimated following equation (1).

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

201   where q_l is the speed of lateral groundwater divergence flux (mm s^{-1}). K_s is the saturated lateral hydraulic

202   conductivity (mm s^{-1}). Thus, the groundwater flow rate per unit width of the groundwater is given by $q_l N$, where N

203   is the thickness of the groundwater layer perpendicular to the impermeable bedrock (m). Inserting the flux equation

204   (1) into the mass continuity equation yields the Boussinesq groundwater equation (2).

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

206   where f is the effective porosity of the groundwater (m^3 m^{-3}). As implied by the Boussinesq equation, the

207   groundwater properties are considered homogeneous across tiles, including K_s, f, and the bedrock slope θ. $\frac{\partial N}{\partial x}$

208   denotes the groundwater hydraulic gradient between adjacent tiles. At each tile, the water table (H) is determined by

209   the balance of the soil–groundwater and the lateral groundwater fluxes (Equation 3).

210     $$H_i^{j+1} = H_i^j + \left(\frac{r_i^j - {q_l}_i^j}{\rho f} \right) \Delta t \qquad (i = 1, \ldots, n_{HB}) \tag{3}$$

211   where H_i^j is the hydraulic head of the water table (m) at i th height band (HB) at j th time step. Based on the

212   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

213   catchment. ρ is the liquid water density (1,000 kg/m^3) and r_i^j is the liquid water flux between the soil–bedrock

214   column and the water table (mm s^{-1}) at i th height band at j th time step. The physical time step Δt of the model is

215   1,800 seconds (30 minutes). $q_{l_i}^j$ denotes the divergence of lateral groundwater flux at i^{th} height band at j^{th} time step

216   and is set to the difference between the lateral divergence from/to HBs immediately above and below. Particularly,

217   in the HB$_1$ (i.e., the nearest height band to the river reach), $q_{l_1}^j$ is determined by the balance of the groundwater

218   discharge from HB$_2$ and the baseflow (if gaining reach) or the channel infiltration (if losing reach). If multiple tiles

219   exist in an HB, r_i is effectively calculated by arithmetically averaging the r values from each tile belonging to the ith

220   height band.

221   LM4-SHARC also resolves the reach-scale streamflow dynamics, and the resulting hydrograph is used as

222   the time-dependent boundary condition at the interface between the stream and hillslope. The Saint-Venant

223   continuity equation with kinematic wave assumption is solved for the streamflow dynamics. For a kinematic wave,

224   the momentum equation assumes that the energy grade line is parallel to the streambed (De St Venant, 1871;

225   Strelkoff, 1970). Thus, the friction slope (S_f) is assumed to be equal to the streambed slope (S_o), physically implying

226   the flow is primarily governed by gravity and friction (Equation 4). Manning's equation was used to express the

227   flow resistance as implied in Equation 4. Using Manning's equation, the stream discharge Q (m^3 s^{-1}) can be related

228   to the cross-section area of flow U (m^2) using a coefficient that depends on channel roughness, slope, and geometry

229   (α) and an exponent determined by the flow characteristics (β) (Equation 5) (Manning et al., 1890). Given that at

230   any cross section, Q and U are functionally related as Equation 5 so that the continuity equation can expressed as

231   Equation 6.

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

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

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

235   where $\frac{dQ}{dU}$ is the kinematic wave celerity (m s^{-1}), and y denotes the river flow direction coordinate. H_1 is the vertical

236   thickness of the groundwater at HB$_1$, which effectively defines the wetted perimeter at the stream-hillslope interface

237   (m). $\frac{2q_l H_1}{\rho}$ is the lateral inflow per unit channel width (m^2 s^{-1}) as q_l is estimated from Equation 1. Once the spatial

238   derivative of the stream discharge Q (i.e., $\frac{dQ}{dy}$) is resolved by Equation 6, the reach outflow (i.e., discharge at the

239   catchment outlet) was used to inversely estimate the river stage, which was in turn used to determine the lateral

240   hydraulic gradients between the river stage and the water table in HB$_1$.

241   We changed the soil column's bottom boundary condition from zero-flux to a variable-flux boundary

242   condition so as to allow for the two-way interaction between the soil columns and the newly introduced groundwater

243   domain. As shown in Figure 1, due to the hydraulic gradient between the bottommost soil layer and the water table,

244   the vertical liquid flux (r in Equation 3) defines the soil columns' bottom boundary condition if the water table is

245   deeper than 10 m from the land surface (Equation 7 and 8). However, if the water table is within 10 m from the

246   ground, the vertical liquid flux at the soil base is considered zero since the equipotential line is assumed to be

247   vertical in the saturated zone, so the vertical hydraulic gradient is zero (Equation 7). The chosen variable-flux

248   boundary condition allows the soil bottom drainage (SBD) at 10 m depth. This enables the consideration of the

249   effects of groundwater on the unsaturated soil processes, depending on groundwater configuration, such as

250   groundwater properties and the water table depth.

251
$$r_i^j = \begin{cases} \dfrac{K_{stog,i}^j \left(\psi_{btm,i}^j + \Delta L_i^j \right)}{\Delta L_i^j} & (water\ table\ depth > 10\ m) \\[2mm] 0 & (water\ table\ depth \leq 10\ m) \end{cases} \tag{7}$$

252
$$\Delta L_i^j = e_{btm,i} - \left(H_i^j + hl_i \tan\theta \right) \tag{8}$$

253   where r_i^j is the vertical liquid water flux (mm s⁻¹) and the $K_{stog,i}^j$ is the hydraulic conductivity between the

254   unsaturated soil bottommost layer and the water table at i th height band at j th time step (mm s⁻¹), which is calculated

255   by the harmonic mean of hydraulic conductivity values in the bottommost soil layer and groundwater. $\psi_{btm,i}^j$ is the

256   soil matric potential at i th height band at j th time step (m), $e_{btm,i}$ the elevation of the central node in the bottommost

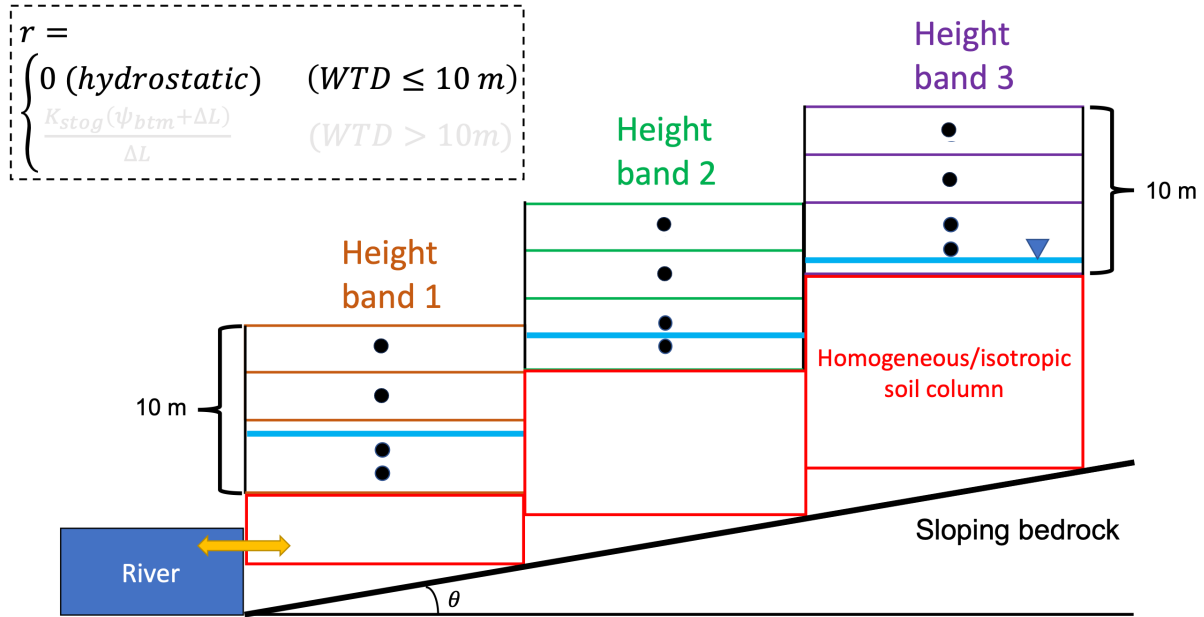
257   soil layer (m). wt_i^j is the pressure head of the water table at i th height band at j th time step (m). hl_i is the total

258   hillslope length from the reach of i th height band characterized by the HMC approach (m). Therefore, Equation 8

259   shows ΔL_i^j is the distance between the soil bottommost node and the water table at i th height band at j th time step

260   (m) considering the bedrock slope θ.

261

(a)

(b)

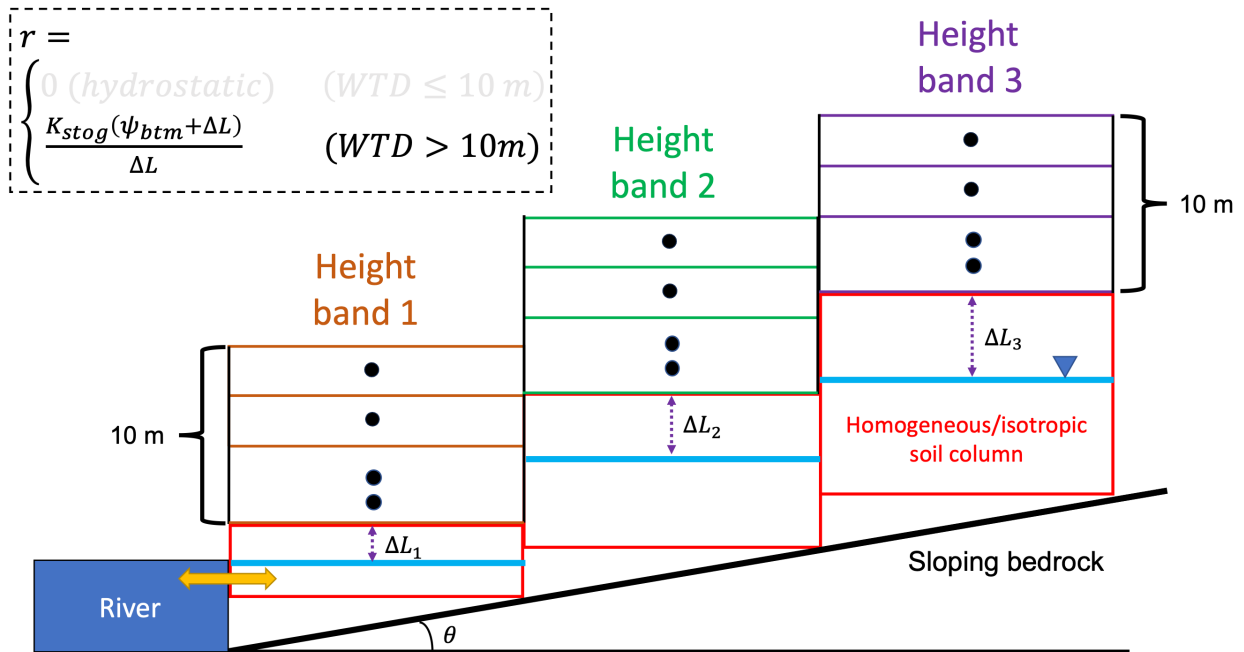



Figure 1. The LM4-SHARC's modified boundary condition (BC) at the soil base depending on the depth of water table. (a) the equipotential line assumed to be vertical (i.e., hydrostatic) if the depth of water table is less than 10 m, (b) the soil base BC changes from zero-flux to variable-flux according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

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 (Equation 9).

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

where wl_{gw}, ws_{gw} are the liquid and ice content in 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 $3.3358 \cdot 10^{5}$ J kg^{-1}. hc_{gw} is the dynamic heat capacity of the groundwater (J K^{-1} m^{-2}). The groundwater temperature in each HB 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 (Equations 10).

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

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 (J m^{-2} s^{-1}). The equations of the vertical heat conduction ($\delta_{v_{cnd,i}}^{j}$) and advection ($\delta_{v_{adv,i}}^{j}$) based on dynamic groundwater heat capacity are addressed in the Appendix A. The direction for $\delta_{v_{adv,i}}^{j}$ 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 the LM4-SHARC. The lateral heat transport in the groundwater domain ($\delta_{l_i}^{j}$) is another component in determining the groundwater temperature (Equations 11, 12, and 13).

$$\delta_{l_i}^{j} = -q_{l,i}^{j}\,clw\left(T_{gw,i}^{j} - T_{freeze}\right) \qquad\qquad (i = n_{HB}) \tag{11}$$

$$\delta_{l_i}^{j} = clw\left[q_{l,i+1}^{j}\left(T_{gw,i+1}^{j} - T_{freeze}\right) - q_{l,i}^{j}\left(T_{gw,i}^{j} - T_{freeze}\right)\right] \qquad (i = 2,..,n_{HB} - 1) \tag{12}$$

$$\delta_{l_1}^{j} = \begin{cases} clw\left[q_{l,2}^{j}\left(T_{gw,2}^{j} - T_{freeze}\right) - q_{l,1}^{j}\left(T_{gw,1}^{j} - T_{freeze}\right)\right] & (q_l \geq 0) \\ clw\left[q_{l,2}^{j}\left(T_{gw,2}^{j} - T_{freeze}\right) - q_{l,1}^{j}\left(T_{stream}^{j} - T_{freeze}\right)\right] & (q_l < 0) \end{cases} \tag{13}$$

293   where $\delta_{l_i}^{\ j}$ is the lateral groundwater heat flux (J m^{-2}s^{-1}). The groundwater temperature at i th height band at j th time

294   step ($T_{gw,i}^{j}$) is determined based on the time-dependent heat capacity and the vertical/lateral heat fluxes from/to i th

295   height band at j th time step (Equation 14).

$$T_{gw,i}^{j+1} = T_{gw,i}^{j} - \frac{\left[\left(\delta_{v_{cnd,i}}^{j} - \delta_{v_{adv,i}}^{j}\right) + \delta_{l_i}^{j}\right]\Delta t}{hc_{gw,i}^{j}} \tag{14}$$

297   For the lateral heat exchange fluxes between the stream and HB$_1$, the stream temperature is considered if

298   the channel loses water to the riparian zone (i.e., $q_l < 0$). The stream temperature is estimated considering how

299   much heat flows into the stream from the hillslopes and out of it through the catchment outlet (Equation 15).

$$T_{stream}^{j+1} = T_{stream}^{j} + \frac{\left[\left(\delta_{l_1}^{j} + \delta_{unsat_i}^{j}\right) - clw\left(T_{stream}^{j} - T_{freeze}\right)\frac{\rho Q^j}{A^j}\right]\Delta t}{hc_{stream}^{j}} \tag{15}$$

301   where $\delta_{unsat_i}^{\ j}$ is the heat flux advection from the unsaturated soil to the reach by interflow (J m^{-2} s^{-1}). Q_i^{j} is the

302   stream discharge at the catchment outlet (m^3 s^{-1}), and A^j is the flow area at the outlet at i th height band at j th time

303   step. $hc_{stream,i}^{j}$ is the heat capacity of the catchment outflow according to the streamflow hydrograph at the outlet (J

304   m^{-2} K^{-1}). Consequently, in LM4-SHARC, states and fluxes for each domain are determined by accounting for the

305   two-way water/heat exchanges among the unsaturated soil, the hillslope aquifer, and the river.

306

307   **2.4. The streamflow recession analysis for the groundwater parameterization**

308   Brutsaert and Nieber (1977) showed that the time derivative of the recession hydrograph can be expressed

309   as a function of streamflow Q (Equation 16). Since the analytical solutions to the Boussinesq equation can be recast

310   in the form of a power law, the Boussinesq groundwater can be effectively characterized based on groundwater

311   parameters such as K_s, f, the initial saturated groundwater thickness N_{ini}, and the contributing area A (Hong and

312   Mohanty, 2023a; Rupp and Selker, 2006; Brutsaert and Nieber, 1977; Szilagyi et al., 1998; Tashie et al., 2021; Hong

313   and Mohanty, 2023b).

$$-\frac{dQ}{dt} = aQ^b \tag{16}$$

315   where b is a constant, and a is a function of groundwater properties. Since the geometric similarity of a unit-width

316   Boussinesq groundwater throughout the entire catchment is assumed, the catchment outflow Q was estimated as

317      $Q = 2q_lHL_{reach}$ (where L_{reach} is the length of the reach). This geometric similarity assumption is also reflected in

318   the numerical estimation of the catchment outflow hydrograph (Equation 6). Because the recession parameters a and

319   b are readily estimated by logarithmic regression of Q on $-\frac{dQ}{dt}$, streamflow observations can be used to infer the

320   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 time (i.e., low flow) domains. Since the 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 Equation 17 with b_{early} being set to 3.0, and the parameter a for the late time domain (a_{late}) is defined in Equation 18 with b_{late} being set to 1.0. As a result, Equation 17 and 18 were used to interpret the intercept (i.e., $\log(a)$) and slope (i.e., b) of logarithmic regression of Q on $-\frac{dQ}{dt}$ derived from observational streamflow data in the early time and late time domain, respectively.

$$a_{early} = \frac{1.33}{K_s f N_{ini}^3 L^2 \cos\theta}, \qquad b_{early} = 3.0 \tag{17}$$

$$a_{late} = \frac{\pi^2 p K_s N_{ini} L^2}{f A^2} \cos\theta \left[1 + \left(\frac{\frac{B}{T_{ini}} \tan\theta}{\pi p} \right)^2 \right], \qquad b_{late} = 1.0 \tag{18}$$

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 by $B=A/(2L)$.

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 of recession slope characteristics among individual REs, which results in different estimates of groundwater properties. The onset of the REs was starting 5 days after the peak to exclude the influence of overland flow (i.e., runoff) on streamflow hydrograph. We examined the consecutive decline of 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.

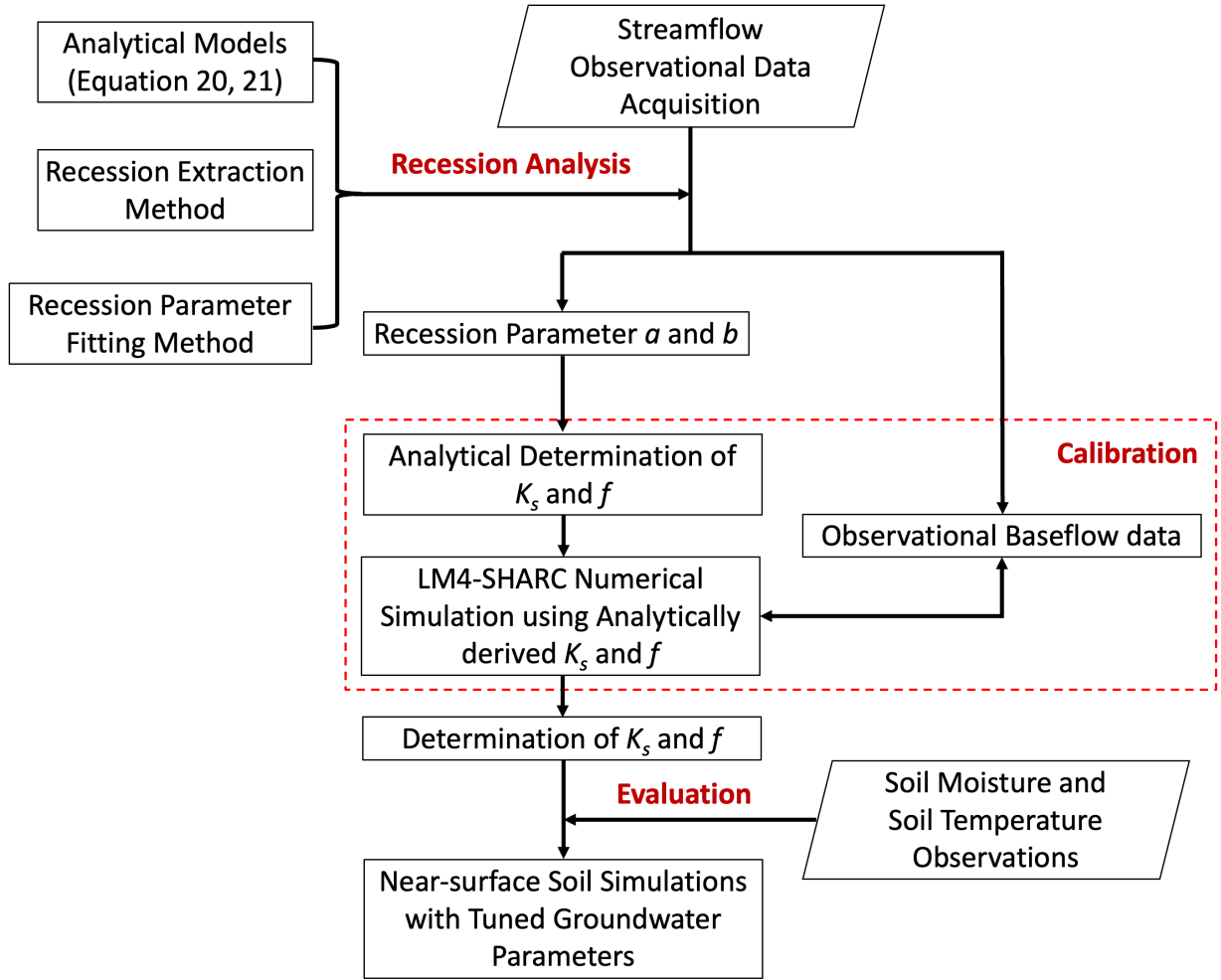
For each RE, we performed the logarithmic regression of Q on $-\frac{dQ}{dt}$ in bi-logarithmic space. The time derivative of Q ($\frac{dQ}{dt}$) was computed based on the daily streamflow. Each RE's transition point of the hydraulic regime from the early to the late time domain was identified daily through 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 an 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 days to get enough data pairs of

353   $\log(Q)$-$\log(-\frac{dQ}{dt})$ for hydraulic regime distinction. For further details on the method for identifying the hydraulic

354   regime transition in an event-scale recession analysis, see Hong and Mohanty (2023a). The process workflow of our

355   method determining the catchment-scale hydraulic diffusivity parameters by combined use of the analytical and

356   numerical models is described in Figure 2.

357

358
359   **Figure 2**. A workflow diagram of the presented method to calibrate and evaluate the catchment-scale groundwater diffusivity
360   parameters based on streamflow recession analysis and combined use of the selected analytical models and LM4-SHARC
361   numerical simulations.

362

363   **2.5. Experimental design for model comparison**

364   LM4-SHARC's new parameterization for the groundwater and its interaction with the soil and river was

365   evaluated based on comparing the respective baseflow and near-surface soil moisture/temperature outputs from the

366   retrospective run of LM4-SHARC and LM4-HydroBlocks. For spin-up, we used periodically cycled GSWP3 10-

367   year forcing. The GSWP3 forcing from 1901-1910 was used repeatedly in each model cycle until a steady-state in

368   the groundwater-related model variables by replacing the initial condition for a new spin-up with the final output

369   state from the previous cycle. The groundwater-related variables include 1) SMC at the bottommost soil layer, 2)

370 baseflow, and 3) water table, and we considered the model to reach a steady state if the simulated difference

371 between the end of the nth and $(n-1)^{th}$ cycle for the respective variables satisfied our criteria simultaneously. In this

372 study, we set the criteria for each variable as 0.001 m^3 m^{-3} (0.1 %) for the bottommost layer's SMC, 0.1 m d^{-1} for

373 baseflow, and 0.001 m for water table. We note that soil-groundwater two-way fluxes (r in Equation 7) was

374 additionally considered in LM4-SHARC for evaluating a steady-state as it is a new variable in LM4-SHARC.

375   Using the confirmed steady state outputs as an initial condition, both model configurations were compared

376 for the period from 10/1/2003 – 9/30/2014 (WY2004 – 2014, 11 years). In this study, the in-situ precipitation and air

377 temperature were assimilated into the GSWP3 forcing data by being directly inserted (i.e., direct insertion data

378 assimilation). This was to improve the accuracy of the model outputs, given the inconsistency between the GSWP3

379 forcing and the local meteorological conditions primarily due to scale difference (i.e., 0.5°×0.5° vs. 1 km^2). We

380 considered precipitation and air temperature the most significant atmospheric variables determining the catchment's

381 water and energy budgets, so that the in-situ precipitation and air temperature data obtained during the evaluation

382 period replaced the corresponding variables in the GSWP3 forcing. LM4-HydroBlocks and the LM4-SHARC were

383 then operated using the identical forcing input. Two statistical metrics of Pearson's correlation coefficient R^2 and

384 *RMSE* (Root Mean Square Error) were used to evaluate the temporal agreement of the modeled hydrologic outputs

385 against corresponding observations and the errors, respectively.

386

387 **3. Study area and observational data**

388 **3.1. Study area and the HMC parameters**

389   The study area is the 1 km^2 Providence Creek P301 headwater catchment in the Sierra National Forest,

390 Nevada (Figure 3 a). The P301 headwater catchment is one of eight primary headwater catchments of the Kings

391 River Experimental Watershed (KREW) project. The U.S. Department of Agriculture's (USDA) Pacific Southwest

392 Research Station initiated and operated the KREW project, part of the National Science Foundation's Southern

393 Sierra Critical Zone Observatory (Jepsen et al., 2016; Hunsaker et al., 2012). The eight catchments are clustered into

394 two groups, Providence and Bull, and, in this study, the P301 headwater catchment that belongs to Providence Creek

395 was selected. We selected the P301 catchment because it contains only one first-order reach with no tributaries into

396 the reach. Since the connectivity between the catchment and reach was required to be established for developing the

397 LM4-SHARC model, the hydrologic configuration of the P301 catchment was considered ideal for evaluating the

398 developed LM4-SHARC model.

399   Surface elevations in the P301 catchment range from 1755 to 2114 m (Hunsaker et al., 2012), on an

400 average topographic slope of 19° The length of the first-order reach in the P301 catchment is 1.5 km, and the

401 average width of the reach was approximated at 10 m to define the channel geometry. The Providence P301

402 catchment represents a rain-snow mixed-conifer forest site with an annual mean precipitation of 1,315 mm/year. The

403 site has a Mediterranean climate, with cool, wet winters and dry summers from approximately May through October

404 (Safeeq and Hunsaker, 2016). During the WY 2004-2014 evaluation period, about 90 % of precipitation occurred

405 between November and June. The mean annual air temperature was measured at 6.8°C. Precipitation falls as a mix

406 of rain and snow, and precipitation transitions from majority rainfall to majority snow approximately at 2000 m in
407 elevation (Bales et al., 2018; Hunsaker et al., 2012).
408   In compiling the land input dataset for the given catchment, the k was set to 1 as the study area is a single
409 catchment. The surface elevation data from 3DEP DEM was used as the sole variable to account for the intra-
410 catchment variability of terrain properties, and each HB's spatial extent was determined using a dh value of 20 m. p
411 was set to 1; however, we note that the number of tiles in each HB can increase (or decrease) depending on factors
412 such as natural mortality, land use, and fire events applied to each tile. The stream flowline was also delineated
413 according to the 3DEP DEM using the area threshold of 100,000 m^2. Consequently, the P301 catchment was
414 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
415 field measurements (Figures 3 b, c, and d).
416
417
418

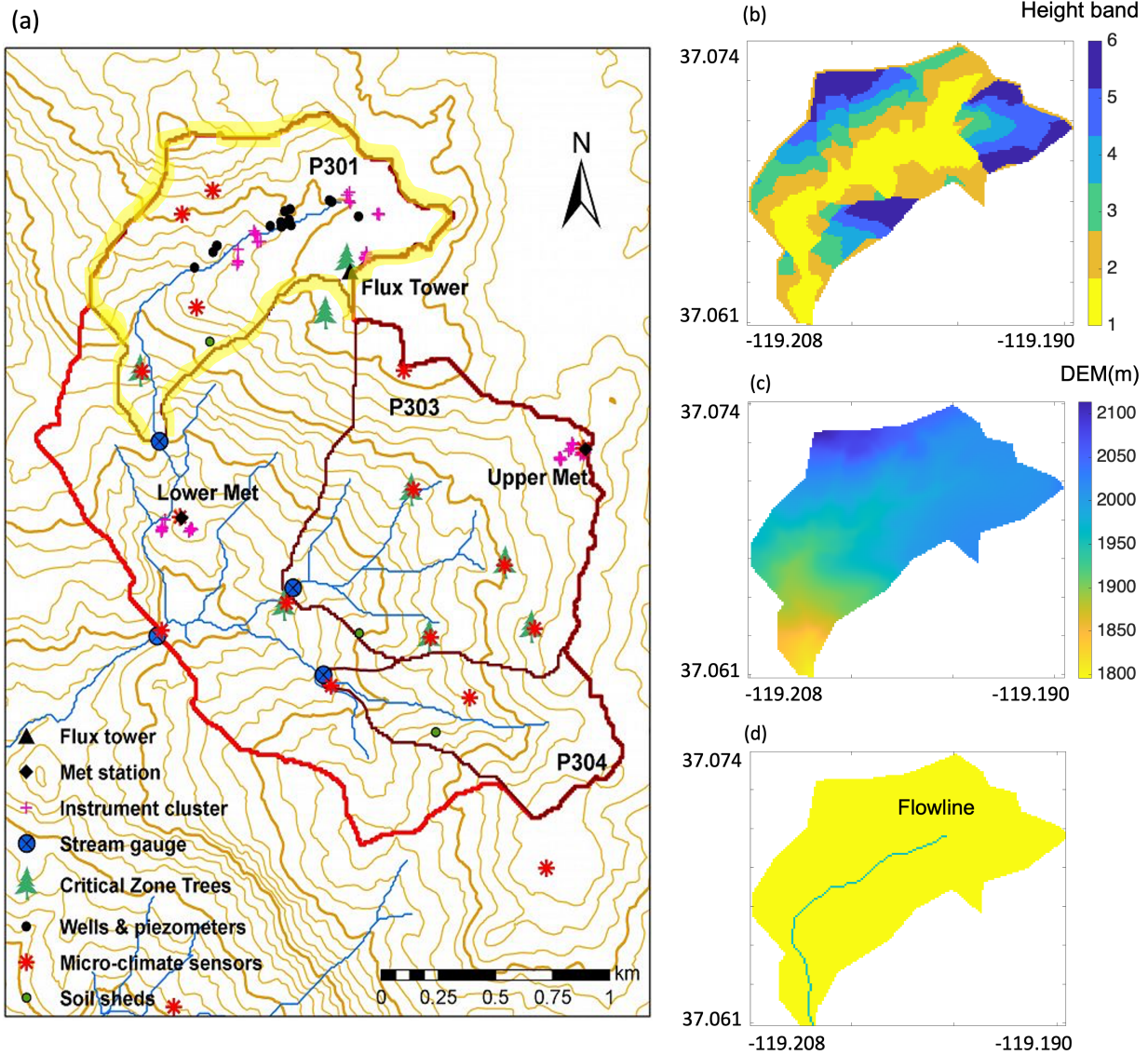



Figure 3. (a) The yellow highlighted area indicate the spatial extent of the headwater catchment P301. While the soil moisture and temperature data measured at the Lower Met station is not within the P301 catchment, the measurements from the Lower Met station was used due to its proximity to the catchment and superior quantity and continuity of observational data compared to other measurement points within the P301 catchment, (b) By 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 pre-processing. Further division of the study catchment into hillslopes is indicated in Supplemental Information S.1.

3.2. Observations for models evaluation

3.2.1. Streamflow and baseflow

The 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-minute intervals and converted to the discharge rate using the standard rating curve supplied by the flume and weir manufacturers (Bales et al., 2018). The

434 stream discharge monitoring began in September 2003 (i.e., WY 2004). The streamflow was averaged daily from

435 WY 2004 to WY 2014 (i.e., 10/01/2003 – 09/30/2014, 11 years in total) and used to evaluate the daily-basis

436 simulation outputs in this study. The daily streamflow observational data derived the daily baseflow rate using the

437 baseflow separation method suggested by Szilagyi and Parlange (1998). The applied baseflow separation method

438 assumes that the drainage from the Boussinesq groundwater maintains the stream recession flow. Thus, the baseflow

439 separation method ensures consistency between the applied analytical models and the observational baseflow data.

440 Szilagyi and Parlange's separation method requires the catchment area, length of channels, and the initial

441 groundwater thickness as parameters. For consistency, the values of these parameters were determined following the

442 parameters used in the analytical models for recession analysis (Equations 17 and 18). Therefore, the catchment area

443 (A) and channel length (L) are considered static according to the field measurement (area: 1 km^2, length: 1.5 km).

444 Applying A value of 1 km^2 assumes that the entire catchment area contributes to the streamflow recession

445 characteristics. However, if only a portion of the catchment contributes to baseflow, this assumption could create

446 uncertainty in the baseflow estimates (Hong and Mohanty, 2023a). The initial saturated groundwater thickness N_{ini}

447 assumed to be constant at a value of 10 m across the individual REs, equivalent to the initial value of groundwater

448 thickness applied in the LM4-SHARC, could also be a source of uncertainty in the baseflow estimates. However,

449 due to the lack of observation-based A and N_{ini} values, this study conducted baseflow separation under these

450 assumptions for the two parameters.

451

452 **3.2.2. Soil moisture, soil temperature, and snow depth**

453 The soil volumetric water content (i.e., soil moisture content (SMC)) and soil temperature (ST) were

454 measured at the Lower Met using ECHO-TM sensors (METER Group). The SMC and ST were measured at 10, 30,

455 60, and 90 cm below the soil surface, and the measurements were used as representative values for each soil depth

456 above and below each sensor(Bales et al., 2018). The measurements at the Lower Met station were used due to its

457 proximity to the P301 headwater catchment (~480 m to the outlet of the P301 catchment). The elevation difference

458 between the Lower Met station and the nearest drainage point (i.e., reach) is about 13 m. At the station, the sensor

459 nodes were installed in locations with different canopy coverage characteristics, such as drip edge, under canopy,

460 and open canopy, to account for the effect of shading (i.e., radiation interception due to canopy) on SMC and ST.

461 Snow depth was also measured at the same Met station, and the distance to snow (or soil if no snow cover exists)

462 was measured using an acoustic depth sensor located at 3 m above the soil surface (Judd Communication LLC)

463 (Bales et al., 2018). The sensors were installed in 2008 (i.e., WY 2009); however, due to the availability of the 10

464 cm depth SMC observations at the open canopy spot, the observation and simulations were compared from WY

465 2009 – WY 2012 (4 years). The SMC and ST observations measured at the depth nearest to the land surface (10 cm)

466 were used to evaluate the near-surface modeled outputs from the LM4-SHARC and LM4-HydroBlocks.

467

468 **3.2.3. Meteorological observations**

469 The meteorological data for model forcing were obtained from a weather station located at the Lower Met

470 station (elevation 1750 m), consistently with the SMC, ST, and snow observations. Precipitation was measured with

471   a Belfort 5-780 shielded weighing rain gauge (Belfort Instrument) located 3 m above the soil surface. The air

472   temperature sensor is 6 m above the soil surface (Vaisala Corporation) (Bales et al., 2018).

473

474   **4. Results and discussion**

475   **4.1. Suitability of the select analytical model**

476   We compared the analytical and numerical solutions of baseflow flux for different combinations of the

477   groundwater diffusivity (D) and bedrock slope (θ) to test their agreements required for applying analytically derived

478   groundwater properties to the numerical groundwater domain. We defined the case of slow groundwater flow

479   (SLW) with K_s of $2 \cdot 10^{-3}$ mm s^{-1} and f of 0.2, K_s of $9 \cdot 10^{-3}$ mm s^{-1} and f of 0.05 for normal groundwater flow

480   (NRM), and K_s of $2 \cdot 10^{-2}$ mm s^{-1} and f of 0.02 for fast groundwater flow (FST). In each diffusivity case, the

481   baseflow was also simulated for distinct bedrock slope conditions, such as a mild bedrock with a tan θ of 0.001

482   degree (MLD) and a steep bedrock with a tan θ of 0.2 degree (STP). The hydraulic conditions, except the D and tan

483   θ, for this gaining reach experiment are: (1) river stages at j th time step at the discharge boundary follow a power

484   function using the initial river stage $h_s^j = h_s^0 \times t^{-0.01}$, representing a falling limb of the hydrograph after peak

485   discharge, (2) the initial head difference between initial saturated groundwater thickness (N_{ini}) and h_s was set to the

486   half of N_{ini} (i.e., $h_s^j|_{j=0} = N_{ini}/2$). The identical hydraulic conditions were applied to both analytical and numerical

487   simulations.

488   The temporal agreement and the total magnitude of groundwater divergence fluxes per unit width (i.e., q_lN)

489   were investigated during the recession duration of 15 days. As shown in Figure 4, q_lN decreases as the hydraulic

490   gradient between the river stage h_s and water table (at height band 1) reduces due to the discharging groundwater.

491   For the representativeness of the time series, the R^2 and $RMSE$ were estimated for the average q_{lat} at each time step,

492   (i.e., $qN_{ave} = \frac{qN_{stp}+qN_{mld}}{2}$). The and $RMSE$ and R^2 were calculated at 0.00088 m^2 s^{-1}, and 0.98, respectively, in the

493   SLW case, 0.0015 m^2 s^{-1} and 0.97 in the NRM case, and 0.0027 m^2 s^{-1}, 0.96 in the FST case. Although the gaps

494   between the analytical and numerical qN_{ave} showed a bit of a drop in the agreements as the groundwater diffusivity

495   increased, we understand the similarly good estimation of the daily cumulative numerical qN_{ave} and temporal

496   agreements (R^2 0.96-0.98) could compensate for the gap and yield groundwater discharge estimates accurate enough

497   to model the streamflow recession. Specifically, care needs to be taken when the presented analytical model is used

498   to tune the numerical Boussinesq groundwater with extremely high D and steeper bedrock (Figure 4 c). Except for

499   such hydraulically extreme cases which is unrealistic in real catchments (i.e., $D > 1.0$ mm s^{-1}), the numerical

500   simulation of the Boussinesq groundwater's discharge with analytically tuned parameters can be justifiable.

501

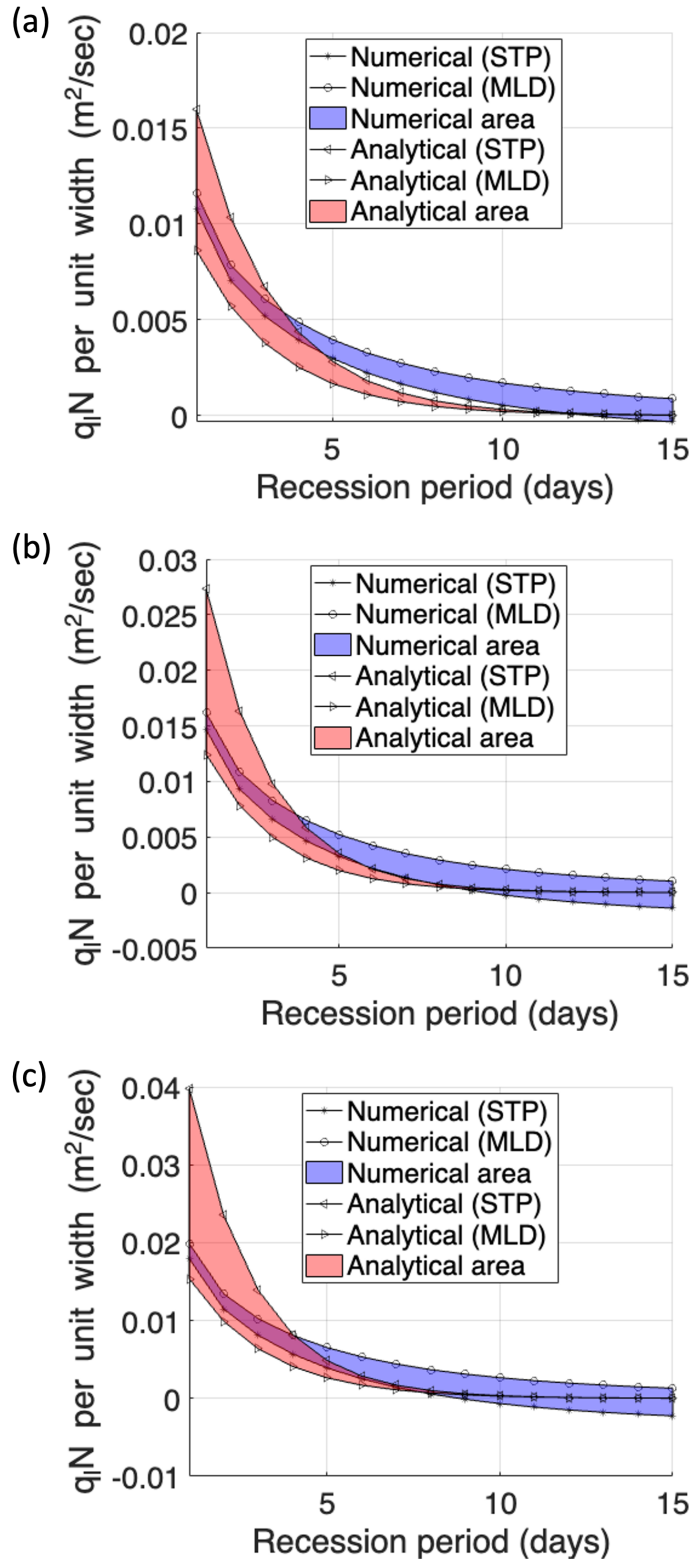



Figure 4. Comparison between the respective baseflow flux solutions derived from the selected analytical and numerical models.
(a) the SLW case with K_s of 2.0E-03 mm/sec and f of 0.20, (b) the NRM case with K_s of 9.0E-03 mm/sec and f of 0.05, and (c)
the FST case with K_s of 2.0E-02 mm/sec and f of 0.02. The numerical and analytical area refer to the uncertainty band of
numerical and analytical solutions due to slope variations, respectively.

4.2. Event-scale recession analysis and calibration

4.2.1. Quantifying the uncertainty of hydraulic diffusivity estimates

We performed the event-scale recession analysis to estimate the effective groundwater properties of the study catchment. Following the extraction criteria (section 2.4.2), 18 individual REs were extracted from the streamflow observations of 13 years, and their recession parameters a and b were estimated for each hydraulic regime (Table 1). Figure 5 (a) presents an example showing how an individual RE was analyzed with the selected analytical models for its early-time 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 hydraulic diffusivity parameters K_s and f. The characteristic length B is calculated by assuming the geometric similarity between both hillslope sides in the catchment; thus, $B=A/(2L)$. The following criteria were applied to define the upper and lower bound 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; Hong and Mohanty, 2023a; 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 of 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}) (Figure 5 b). Essentially, this parameter provides insights into the variability of 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 of 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., $\log\left(-\frac{dQ}{dt}\right) = -3.10\log(dQ) - 6.81$) and the highest a (i.e., $\log\left(-\frac{dQ}{dt}\right) = -2.43\log(dQ) - 6.20$). The upper and lower bounds of K_s were thus estimated as [0.0026 mm s^{-1}, 0.0138 mm s^{-1}]. For f the bounds were identified as [0.033, 0.190], and for θ were [5.0, 17.0 degrees]. Also, for each set of data pairs of $\log\left(-\frac{dQ}{dt}\right)$-$\log(dQ)$ with the lowest and highest a, the relationship between hydraulic diffusivity (per unit wetted perimeter) D and the bedrock slope θ follows a power function. Consequently, we considered that the (θ, D) space in which solutions of D and θ (potentially representing the catchment average behavior) could exist was further constrained by the specified range of parameters K_s, f, and θ between the power functions (Figure 5 b).

RA	Recession period (duration day)	log(a_{early})	log(b_{early})	log(a_{late})	log(b_{late})
1	4/6/04 – 5/12/04 (37 days)	-3.98		-6.95	
2	7/2/04 – 7/28/04 (27 days)	-3.72		-6.76	
3	5/20/05 – 6/8/05 (20 days)	-3.61		-7.10	
4	6/18/05 – 8/14/05 (58 days)	-3.55		-6.75	
5	5/23/06 – 6/12/06 (21 days)	-2.71		-6.84	
6	6/14/06 – 7/17/06 (34 days)	-2.99		-6.33	
7	7/30/06 – 9/6/06 (39 days)	-3.01		-7.66	
8	5/5/07 – 5/30/07 (26 days)	-2.75		-5.72	
9	6/7/07 – 7/7/07 (31 days)	-2.12		-6.76	
			3.0		1.0
10	6/5/08 – 7/12/08 (38 days)	-1.88		-6.45	
11	7/14/08 – 8/18/08 (36 days)	-2.75		-5.98	
12	5/3/09 – 5/29/09 (27 days)	-2.12		-6.60	
13	7/1/09 – 8/5/09 (36 days)	-2.84		-5.61	
14	6/6/10 – 8/1/10 (57 days)	-1.92		-6.67	
15	8/3/10 – 8/27/10 (25 days)	-1.99		-6.11	
16	8/1/11 – 8/22/11 (22 days)	-2.86		-6.32	
17	4/27/12 – 5/24/12 (28 days)	-2.33		-7.12	
18	5/22/14 – 6/10/14 (20 days)	-2.70		-5.49	

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

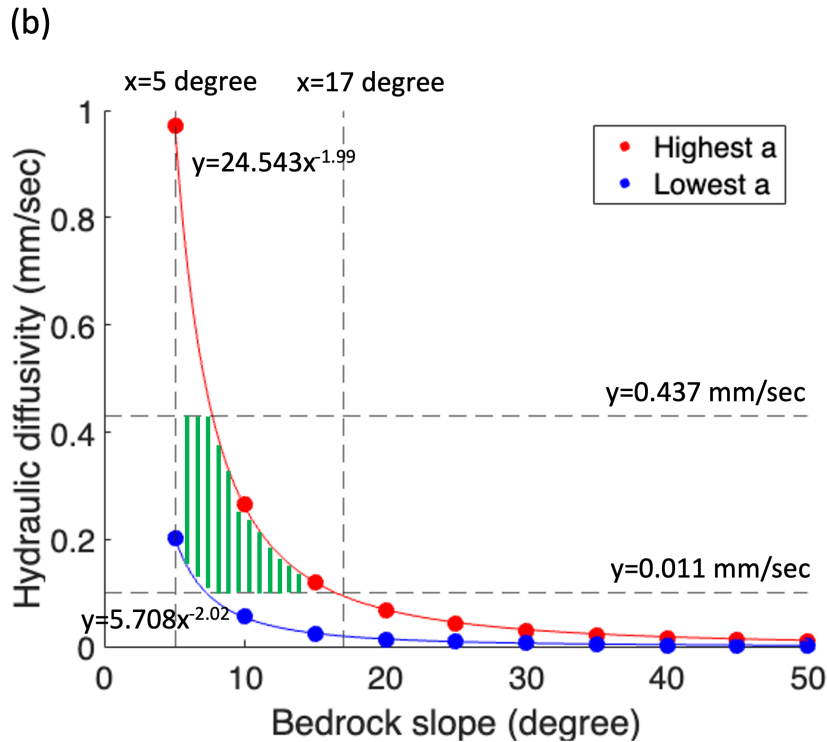
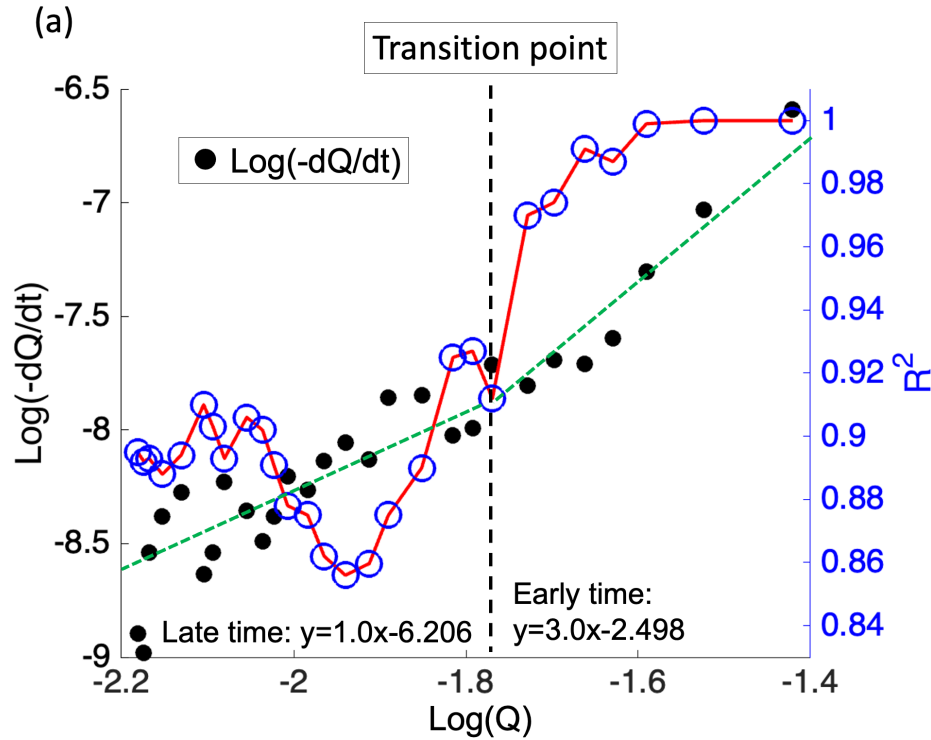
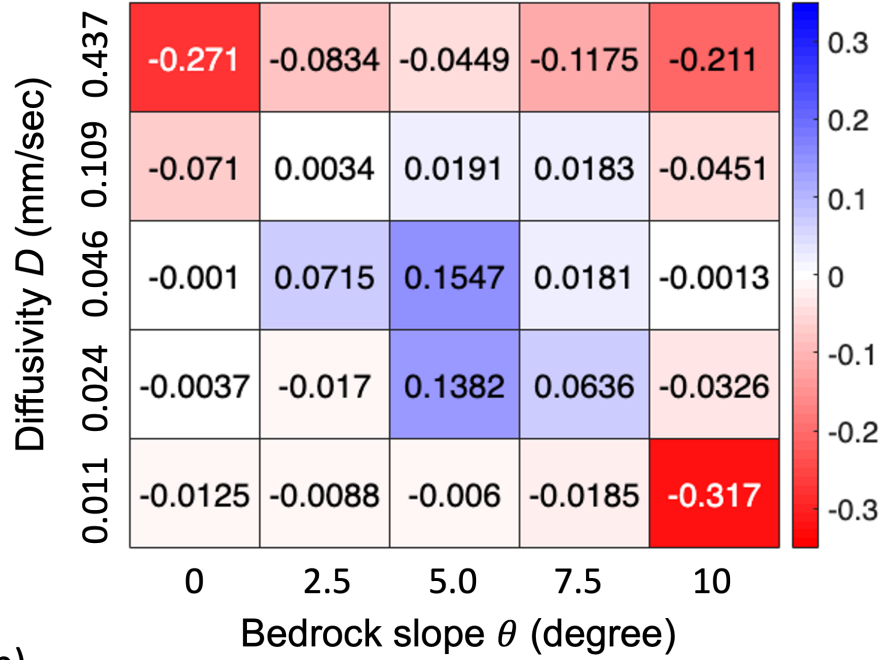



Figure 5. (a) An example showing how the transition point of hydraulic regime (from early-time to late-time domain) is determined from an individual recession event. The combined understanding of the recession parameter a_{early}, and a_{late} enables inferring the groundwater properties such as hydraulic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

4.2.2. Calibrating groundwater properties based on baseflow flux accuracy

While it is known that 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's groundwater domain. Based on the identified value range of θ and D, 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 10/1/2003 – 9/30/2014, WY2004 - 2014) were used for calibration. We considered a (θ, D) that best represents the temporal dynamics and magnitudes of baseflow observations as the calibrated (θ, D) for the study catchment.

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

(a)

(b)

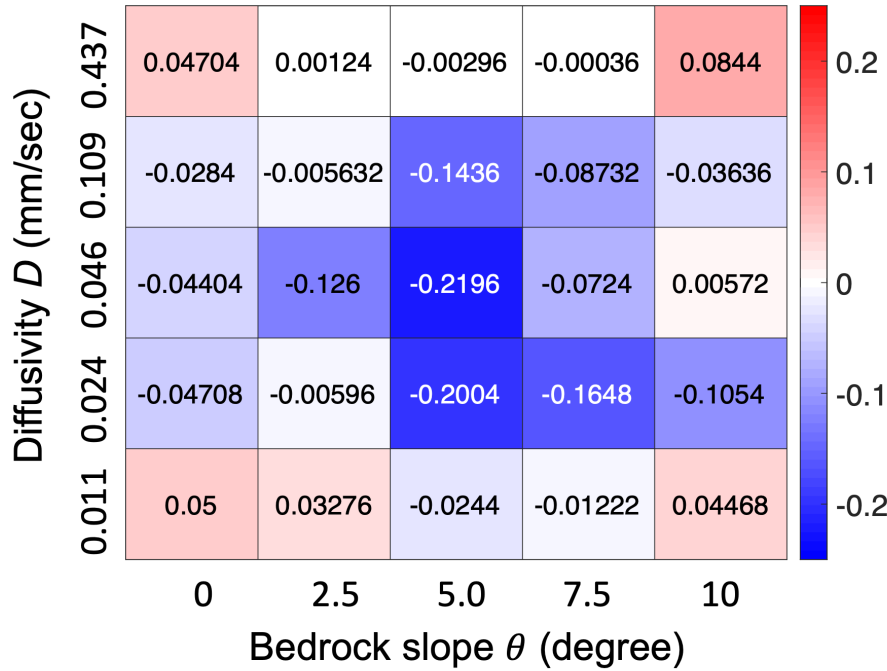



Figure 6. The pair of groundwater parameters (θ, D) was specified to calibrate the catchment groundwater. (a) R^2 difference (improvement) between the LM4-SHARC and LM4-HydroBlocks ($R^2_{LM4\text{-}SHARC}$ - $R^2_{LM4\text{-}HydroBlocks}$). Here, R^2 denotes the coefficient of determination between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

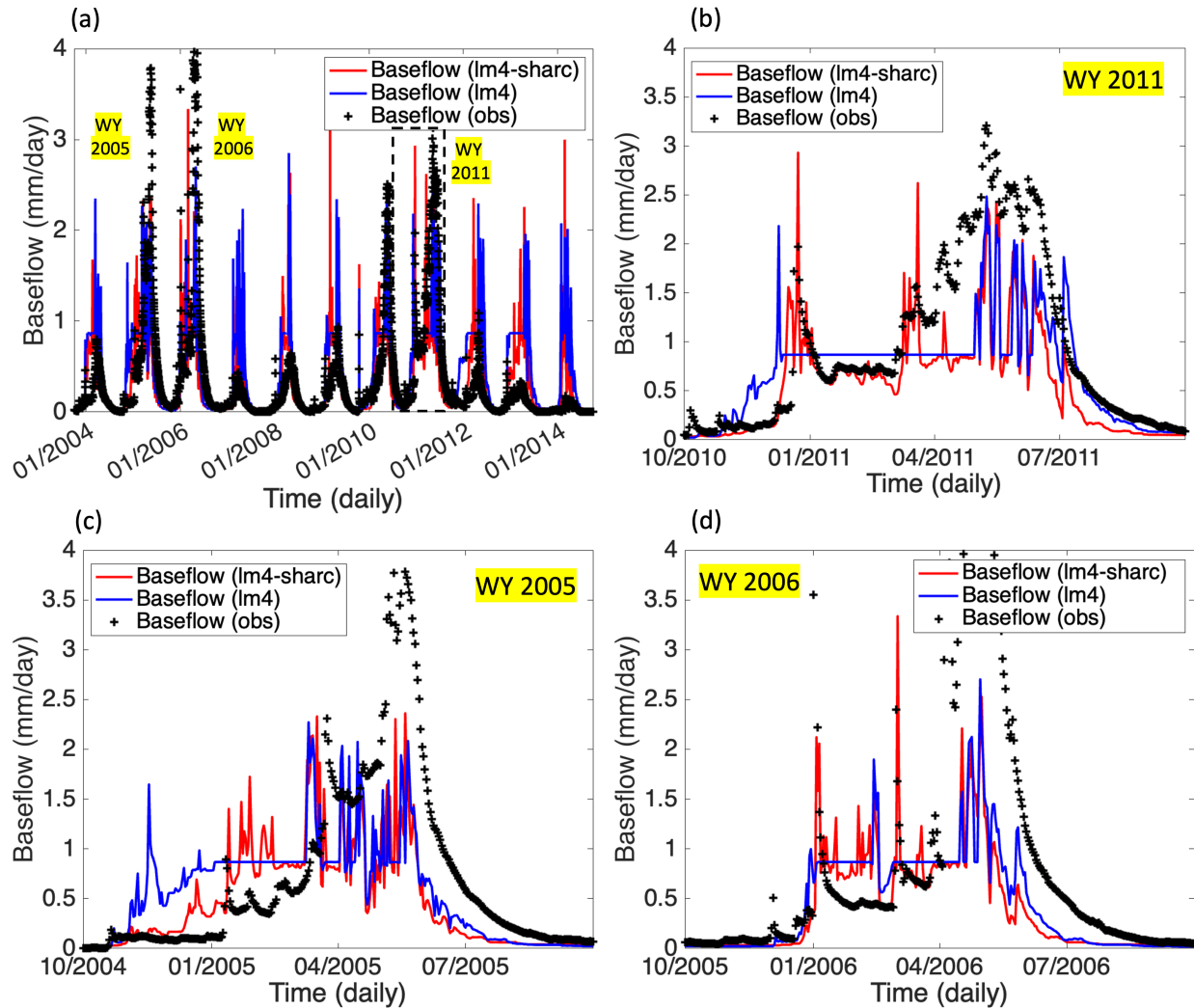



Figure 7. Comparative time-series of baseflow estimates from the LM4-SHARC (red), LM4-HydroBlocks (indicated LM4 in the legend, blue), and the corresponding observations. (a) Daily time-series of the baseflow data over the evaluation period of 11 years, (b) time-series in Water Year (WY) 2011, (c) time-series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

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 θ=5.0 degree, D=4.6· 10^{-2} mm s^{-1} to tune the Boussinesq groundwater domain (in the LM4-SHARC) and compared the respective soil moisture and temperature estimates at 10 cm depth from both configurations. The evaluation of near-surface soil moisture and temperature predictions was performed for four years, from WY 2009-2012. Figures 8 (a), (b), (c), and (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 four years of the evaluation period, we identified that soil bottom drainage (SBD) facilitation due to the modified boundary condition from zero-flux to variable-flux BC at soil columns' base (Figure 1 a) significantly affected the soil moisture content (SMC) at 10 cm depth. The time series of 10 cm SMC from the LM4-SHARC showed significantly reduced SMC compared to LM4-HydroBlocks. The entire soil columns reached full saturation too quickly in the LM4-HydroBlocks, and the facilitated downward liquid transport due to drier soil lower layers in the 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 the LM4-SHARC. The average of the four R^2 from the yearly time series comparison (i.e., 2009-12) was improved from 0.831 (LM4-HydroBlocks) to 0.849 (LM4-SHARC) while noticeably reducing the average $RMSE$ from 0.0722 m^3 m^{-3} (LM4-HydroBlocks) to 0.0425 m^3 m^{-3} (Figure 9 a and b).

We also found that the enhanced representation of SMC resulted in better capturing 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 the decreased SMC (i.e., liquid water) could increase the soil temperature under the given net radiation. Consequently, we identified that soil temperature predictions at the 10 cm depth showed significant improvements in the 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 depth could hold too much soil water due to the imposed shallow water table. This overestimation in soil water content could lead to an inaccurate description of 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 from the yearly time series comparison between the modeled and observed soil temperature increased from 0.952 to 0.957, and the $RMSE$ significantly reduced from 2.77 K to 1.67 K (Figure 9 c 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 melt-out timing of snow involves a mutual influence on soil temperature. Specifically, we noticed that the timing of snow melt-out is affected by soil temperature, as the warmer ground expedites the melting. The melt-out timings of snow in the four evaluation years were represented sooner in the LM4-SHARC than the LM4-HydroBlocks due to increased soil temperature (Figure 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.

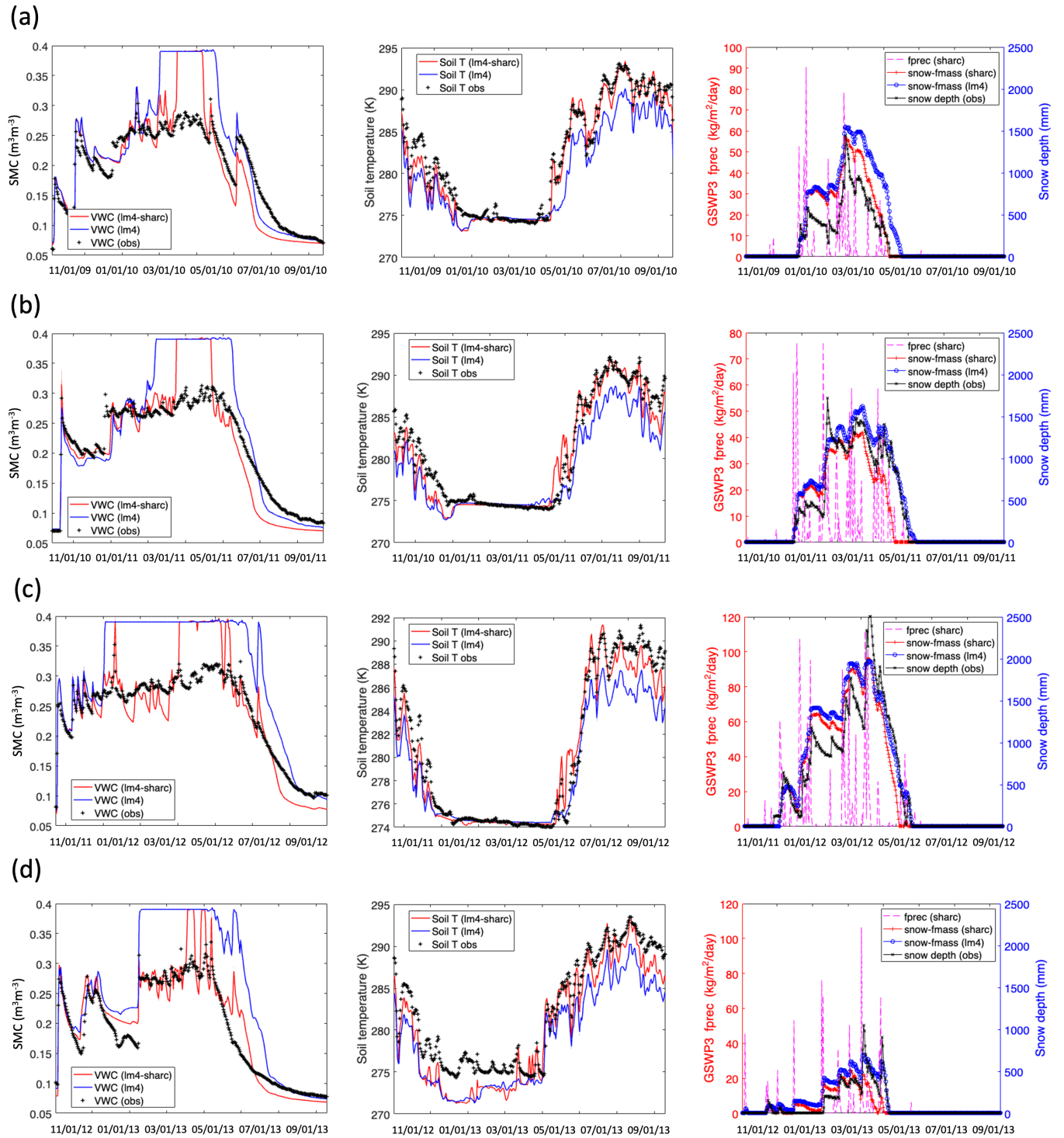



Figure 8. Comparative yearly time-series of soil moisture, temperature, and snow mass for the evaluation period from (a) WY2009, (b) WY2010, (c) WY2011, and (d) WY2012. For all the years, higher soil temperature at 10 cm depth due to drier soil were generally identified. The GSWP3 fprec stands for frozen precipitation, denoting the snowfall rate in unit of kg/m^2/day.

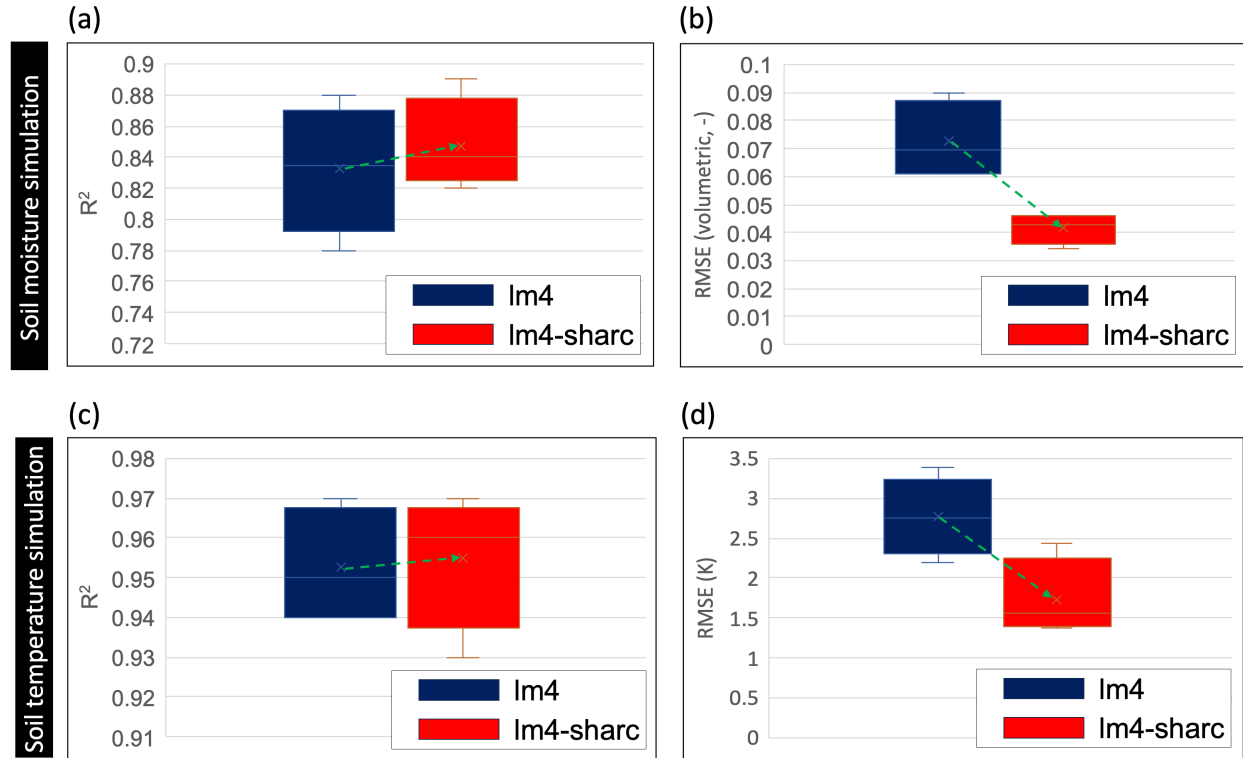



Figure 9. (a), (b) the average of the four R^2 from the yearly time-series comparison of 10 cm depth soil moisture (i.e., 2009-12) was improved from 0.831 (LM4-HydroBlocks) to 0.849 (LM4-SHARC) while significantly reducing the average *RMSE* from 0.0722 m³/m³ (LM4-HydroBlocks) to 0.0425 m³/m³, (c), (d) R^2 from the yearly time series comparison between the modeled and observed soil temperature at 10 cm depth increased from 0.952 to 0.957, and the *RMSE* significantly reduced from 2.77 K to 1.67 K. LM4 denotes LM4-HydroBlocks.

4.3.2. Sensitivity of soil water storage to stream-groundwater diffusivity

After examining the enhancement in 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 the LM4-SHARC and LM4-HydroBlocks. ΔSWS was calculated by subtracting the total SWS per unit area of the soil columns (in the catchment) derived from the LM4-HydroBlocks from that of the 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 of SWS (ΔSWS) is evaluated only after the steady state is reached. Figure 10 (a), (b), (c), (d) show that the simulations of the four variables reduce their deviations (from the previous cycle) as the cycles progress and how the model was considered to be steady-state based on the agreements of variables from the nth cycle and $(n-1)^{th}$. We also identified that the groundwater properties influenced the time to reach a steady state (i.e., spin-up time) as the variations in groundwater diffusivity affect the flux velocity. More detail can be found in Supplemental Information S.2.

ΔSWS 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 the downward

655   hydraulic gradients from the subsurface soil and water table. Figure 11 (a) shows the variability and magnitude of

656   the entire catchment's annual mean of ΔSWS (over ten years) according to θ and D (y-axis values denote average

657   ΔSWS per soil column). It is noticeable that the annual mean of ΔSWS gradually increased negatively (i.e., less

658   SWS in the LM4-SHARC) with increasing D until the water table dynamics did not significantly affect hydraulic

659   gradients between subsurface soil-water table (i.e., lack of groundwater storage). This happened when the downward

660   groundwater recharge fluxes were continuously less than q_l in the groundwater domain. Also, the catchment's total

661   ΔSWS was found to be lower if the slope is steep. While the steeper bedrock showed a greater q_l with the increased

662   gravitationally-driven discharge fluxes, ΔSWS values were more noticeably affected by the lowered water table due

663   to milder bedrock in the groundwater domain. In the case of this catchment, it was also observed that for D values

664   greater than 0.1 mm s^{-1}, the lack of groundwater storage occurred irrespective of the bedrock slope.

665   Moreover, the SBD facilitation due to the groundwater lateral discharge was more noticeable as the height

666   band was farther from the reach. The annual mean of ΔSWS per unit area gradually decreased from HB$_6$ (~120 kg

667   m^{-2} y^{-1}) to HB$_1$ (~10 kg m^{-2} y^{-1}), with a sharp decrease at HB1 (Figure 11 b). It was also noticeable that the distinct D

668   made less difference among the values of ΔSWS per unit area as the height band was farther from the reach. Figure

669   11 (c) and (d) illustrate that the effects of groundwater flow conditions on the SWS variability could be more

670   significant in the riparian/river valley compared to the hilltop area. This also implies that the groundwater's effects

671   on the water content in the partially-saturated soil depend on the depth of water table, leading to higher sensitivity of

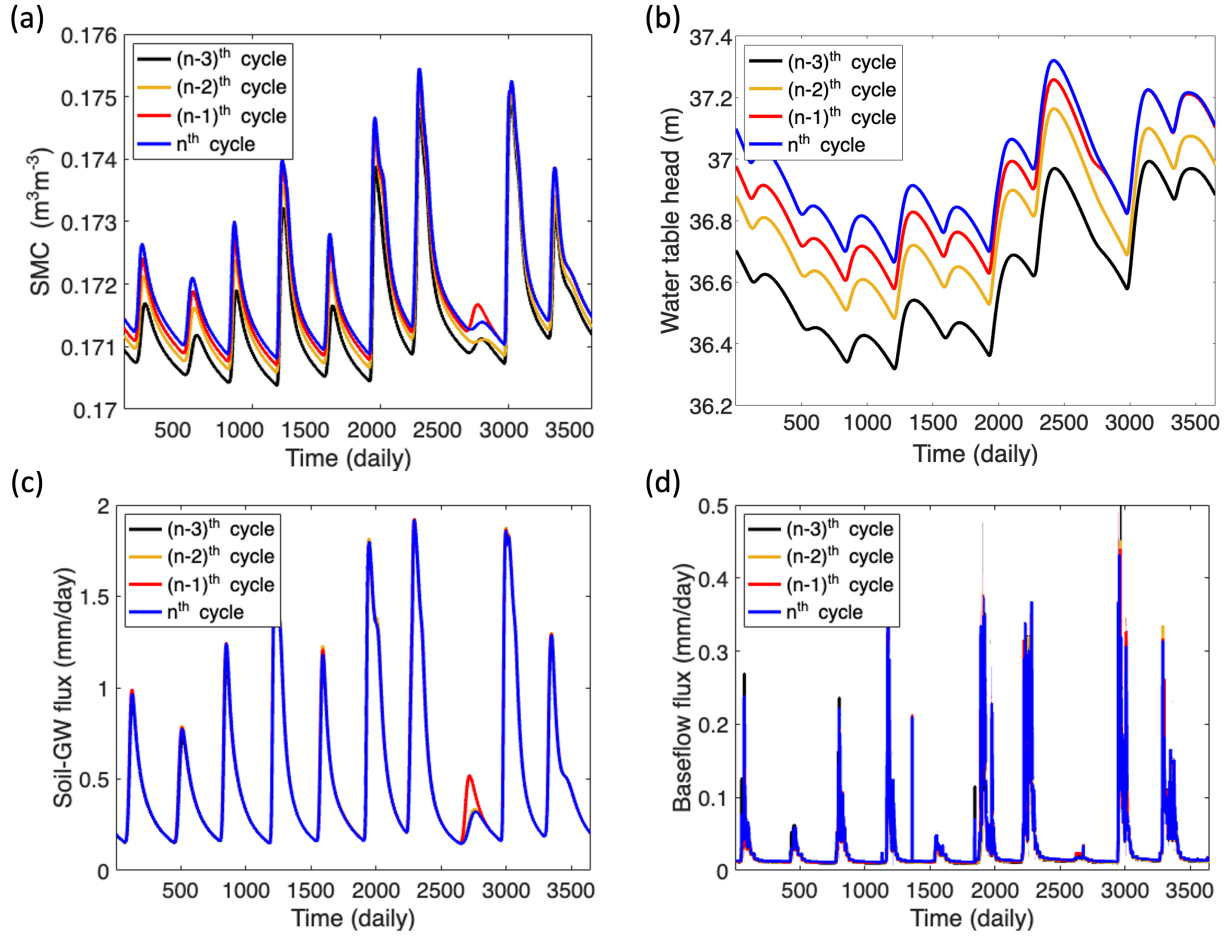
672   SMC to groundwater diffusivity if the water table is shallow than deep groundwater. While the HB$_6$'s annual ΔSWS

673   per unit area (m^2) values were the greatest among the HB$_1$ – HB$_6$, ranging from 121.45 kg m^{-2} y^{-1} (VSW) to 121.89

674   mm (FST), the ΔSWS difference (between VSW and FST) of 0.44 kg m^{-2} y^{-1} was minor compared to the result from

675   the HB1. From HB$_1$, the ΔSWS values were found to range from 7.75 kg m^{-2} y^{-1} (VSW) to 11.70 kg m^{-2} y^{-1} (FST),

676   yielding a difference of 3.94 kg m^{-2} y^{-1} due to the groundwater diffusivity.

677

678
679 **Figure 10**. (a) The volumetric soil moisture content at the soil bottommost layer, (b) the groundwater table, (c) groundwater
680 recharge fluxes (since positive), (d) baseflow flux reduce their deviations (from the previous cycle) as the cycles progress to be
681 considered steady-state based on the differences from the nth cycle and $(n-1)^{th}$.
682
683

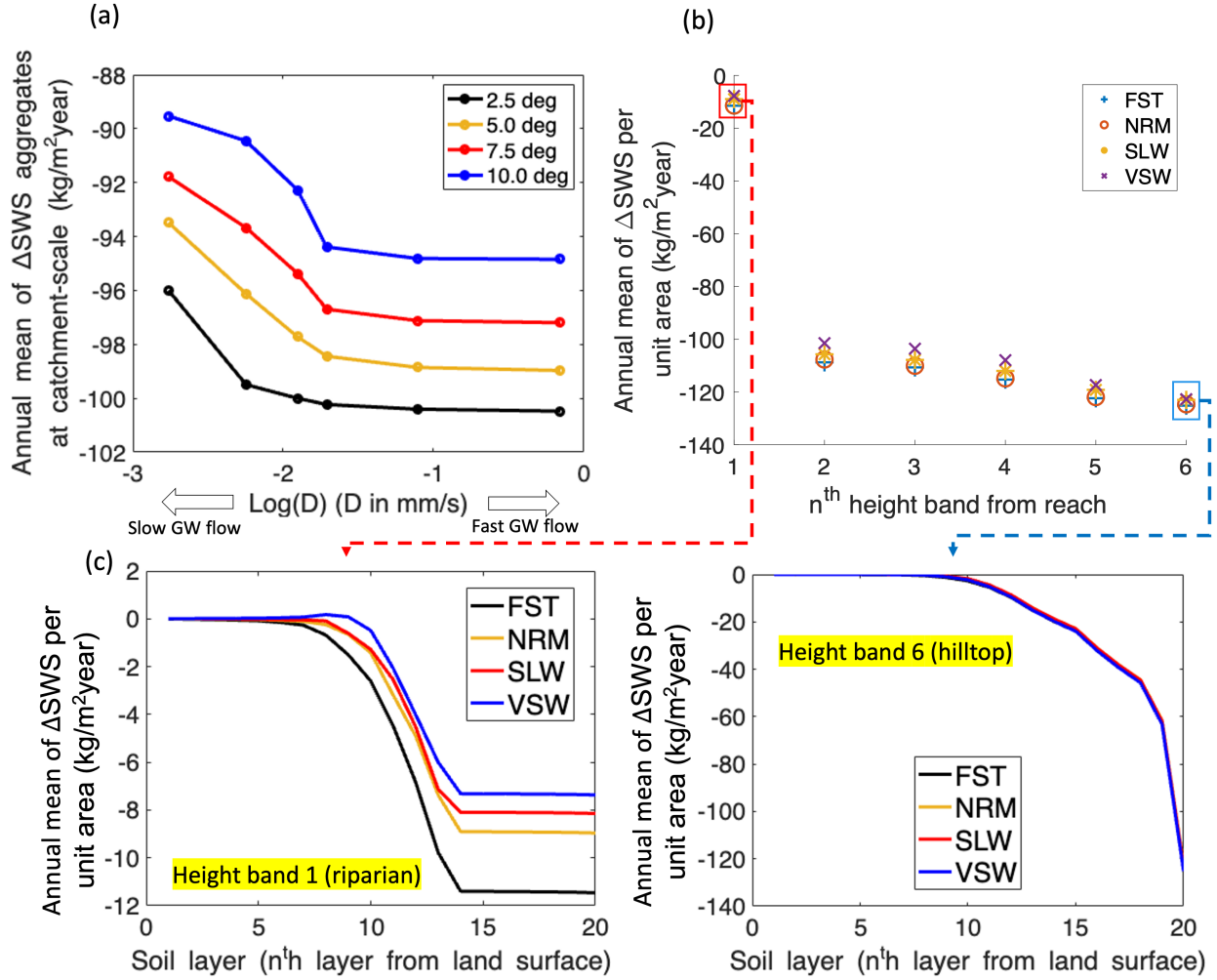



Figure 11. (a) Sensitivity of soil water storage (SWS) to the diffusivity D and bedrock slope θ, (b) the soil drainage (i.e., downward recharge) facilitation due to the groundwater lateral discharge was more noticeable as the height band was farther from the reach, (c), (d) a more significant effects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

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 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 spin-up using the GSWP3 forcing with assimilated in-situ precipitation/air temperature data (i.e., GSWP3 (01/1901-09/2003) + in-situ (10/2003-09/2014)). The comparison of LAI time series between the LM4-SHARC and LM4-HydroBlocks revealed that the differences in the LAI at the hilltop area (i.e., HB_6) were more dramatic than those at the riparian zone (i.e., HB_1) as the vegetation evolved (Figure 12 a and b). With a closer look at the LAI time series for the recent four WYs (i.e., WY 2009-2012 as studied in Section 5.3.1) at the riparian zone and hilltop, moreover,

701   we found that the LAI contrast was more significant during the warmer season at the hilltop area, while the overall

702   trend of the LM4-SHARC-derived LAI evolved comparably to that of the LM4-derived LAI at the riparian zone

703   (Figure 12 c, and d). The hydrologic convergence at the riparian zone explains the comparable LAI dynamics at the

704   riparian zone as the subsoil domain readily saturated by the converging water impedes SBD, leading to higher water

705   retention in the partially saturated soil. Different soil moisture availability (between HB_1 and HB_6) resulting from

706   the contrasting SBD dynamics is thus emphasized by the density of plants, especially during the warm(er) season

707   when the plants yield higher transpiration rates. The varied transpiration rates in the LM4-SHARC also consistently

708   explain the LAI contrast due to different soil moisture conditions. We found that the transpiration rate at the hilltop

709   was reduced by 29.5% in the LM4-SHARC, while the rate was reduced by 10.3% in the riparian zone in the LM4-

710   SHARC (Figure 12 e and f). Overall, we found that the variations in SWS, transpiration, and LAI simulations are

711   consistent in that the groundwater convergence to the river valley intensified the catchment's contrasting hydrologic

712   states.

713

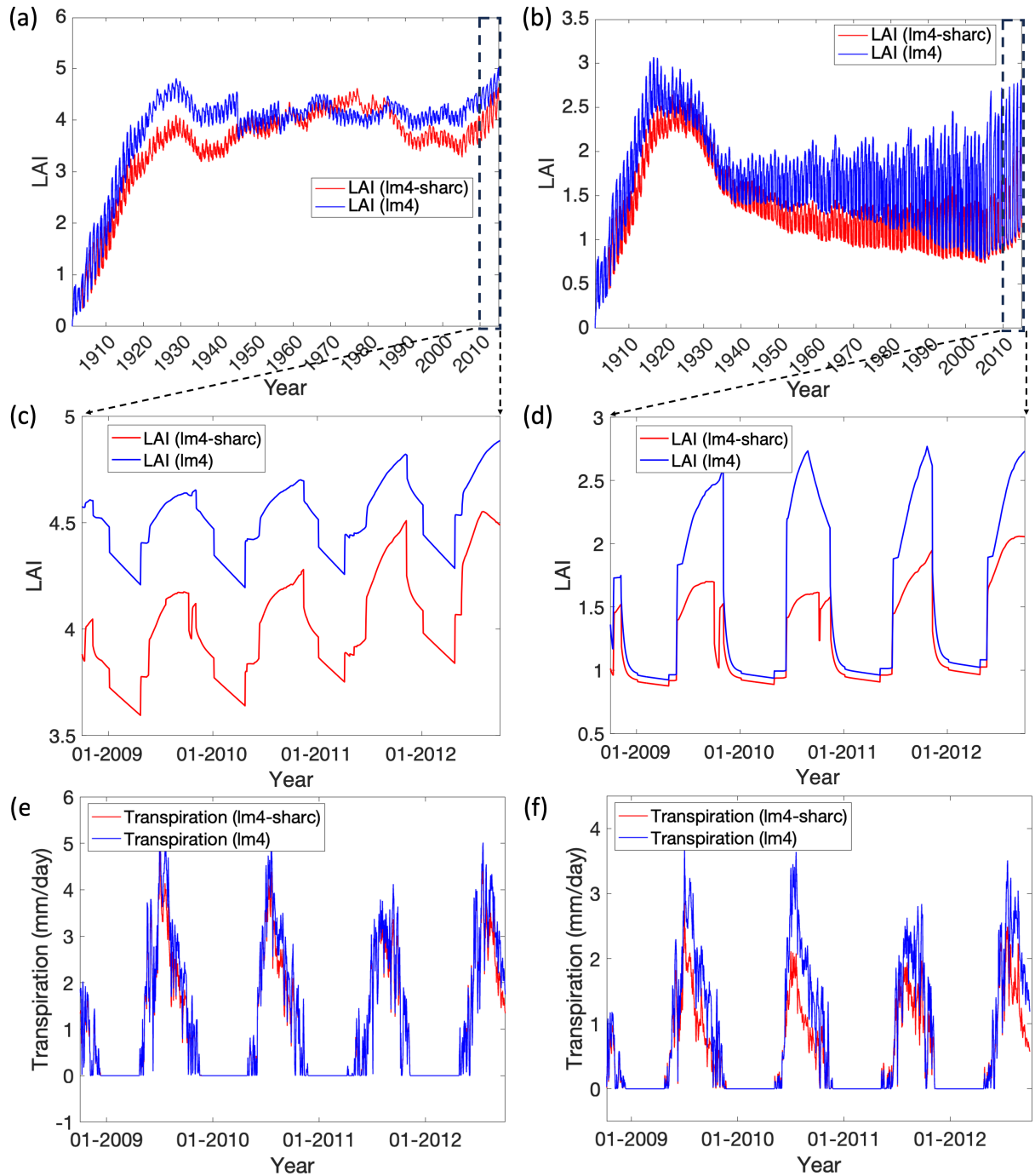



Figure 12. Temporal evolution of LAI from 1901 to 2014 (114 years) using GSWP3 forcing at (a) the riparian zone (HB1) and (b) the hilltop (HB6). The differences in the LAI (between the LM4-SHARC and LM4-HydroBlocks) at the hilltop area were more dramatic than those at the riparian zone. (c), (d) the LAI time series for the recent four WYs (i.e., WY 2009-2012), (e), (f) the time series of plan transpiration rate (mm/day) for the same period. LM4 denotes LM4-HydroBlocks. The effects of varied transpiration dynamics in LM4-SHARC on evapotranspiration estimates are discussed in Supplemental Information S.3.

4.5. Applicability of the 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 (Figure 13 a): 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: 3,136 mm y^{-1}, and steepest: 0.547 m m^{-1} catchment) to those of the Musselshell catchment (the driest: 397 mm y^{-1}, and mildest: 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 that diffusivity D was set to 0.046 mm s^{-1} and the slope of groundwater bedrock θ was assumed to be 5 degrees in all the 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 if the model is robust under diverse conditions by examining the consistency between hydrologic characteristics and indices. The hydrologic indices include: (1) runoff ratio (i.e., ratio of streamflow to precipitation), (2) baseflow coefficient (i.e., ratio of baseflow to precipitation). The annual mean baseflow/streamflow of each catchment was estimated at 10.9 mm y^{-1}/49.8 mm y^{-1}, 105 mm y^{-1}/364.0 mm y^{-1}, and 281 mm y^{-1}/1,897.4 mm y^{-1} from the Musselshell, Maine, and Clearwater catchment, respectively. With the respective annual mean precipitation of each catchment (i.e., Musselshell - 397.2 mm y^{-1}, Maine - 1223.1 mm y^{-1}, and Clearwater - 3136.6 mm y^{-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 Musselshell, Maine, and Clearwater catchment, respectively. The established positive correlation between the baseflow coefficient and runoff ratio is consistent with what is reported in existing studies (Cheng et al., 2021; Ouyang et al., 2018) (Table 2). We also found that the value gradients of baseflow/recharges 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 minimal baseflow amounts due to the lack of groundwater storage from the limited groundwater recharge (Figure 13 b and c).

		Headwater catchment		
		Musselshell (MT)	Maine (ME)	Clearwater (WA)
Hydrologic characteristics	Annual mean precipitation	397.2 mm y^{-1}	1,223.1 mm y^{-1}	3,136.6 mm y^{-1}
	Annual mean streamflow	49.8 mm y^{-1}	364 mm y^{-1}	1,894.7 mm y^{-1}
	Annual mean baseflow	10.9 mm y^{-1}	105 mm y^{-1}	281 mm y^{-1}
	Annual mean groundwater recharge	0.21 mm y^{-1}	150.4 mm y^{-1}	360.2 mm y^{-1}
Indices	Runoff ratio	0.125	0.298	0.604
	Baseflow coefficient	0.028	0.086	0.230

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

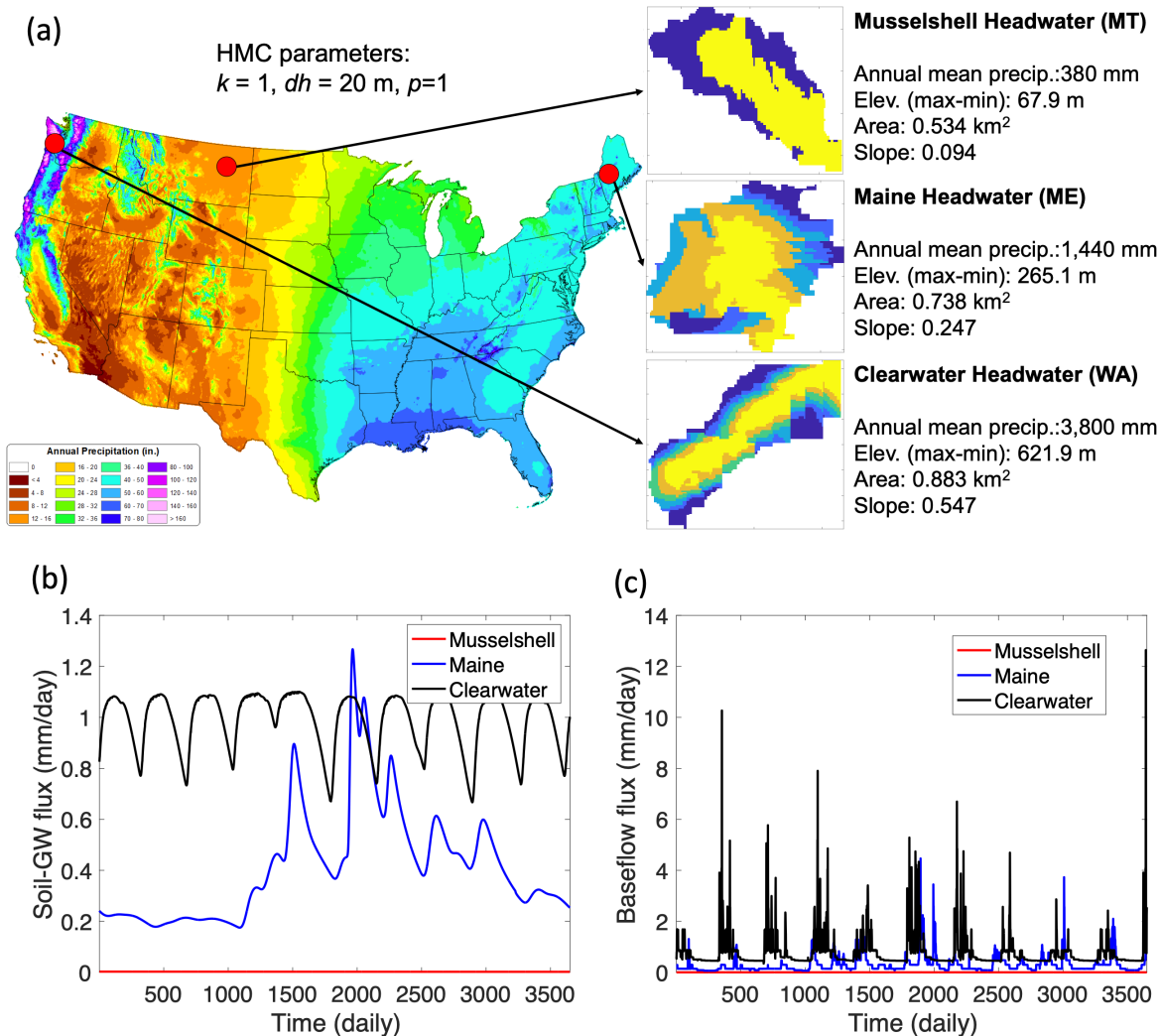



Figure 13. (a) the additional three sites are marked (red dots) on the PRISM 30-year normal precipitation map (2022, PRISM Group, Oregon State University). The sites include Musselshell headwater (MT), Maine headwater (ME), and Clearwater headwater (WA) catchments, (b) the LM4-SHARC's daily soil-groundwater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

4.5.2. Inferring groundwater properties for global-scale SHARC simulations

Since the presented parameterization scheme SHARC relies on the observationally-derived recession characteristics of streamflow, a method must be developed for quantifying the recession variability (using parameters *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, and climatology and remotely-sensed hydrologic data can be used complementarily 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) was established by means of simple regression analysis or machine learning (ML) techniques (e.g., random forest)

(Zecharias and Brutsaert, 1988; Hong and Mohanty, 2023b; Tashie et al., 2021; Cai et al., 2021). Specifically, Hong and Mohanty (2023b) examined the relationship between the local catchment's mean soil permeability, derived from STATSGO2 soil map dataset, and recession parameter a, enabling parameterization of Boussinesq groundwater at a large-scale encompassing 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 the significant relationship between the STATSGO2- derived catchment average soil permeability and observationally-derived recession parameters, and parameterized the catchment-scale groundwater in the entire 33,436 catchments (including 40 gauged and 33,396 ungauged catchments). Furthermore, in addition to global soil properties data, the utility of remotely sensed and global in-situ hydrologic data will be studied to understand if the groundwater recession behavior could be predicted using machine learning (ML) frameworks (e.g., random forest). 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 AirMOSS, 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 for the robustness of groundwater parameters with identified relative importance of each hydrologic variable.

In addition to identifying the relationship between data and recession parameter, directly utilizing baseflow/streamflow estimates available at catchment resolution that enables 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 by assimilating SWOT altimetry data. Based on existing findings about the utility of SWOT data for baseflow estimation, the possibility of obtaining streamflow/baseflow at catchment resolution with high temporal resolution (e.g., daily) will be further explored. Hong and Mohanty (2023a) presented that the NSD data could be key hydrologic variable to infer the Boussinesq aquifer's diffusivity downstream of the river. The NSD data provides information of net mass balance during drought flow by isolating baseflow component using catchments' inflow and outflow data (Rhodes et al., 2017). To this end, we will explore the data-aided streamflow estimates at catchment resolution to use them as surrogate data (of actual streamflow measurements) to establish the catchment-scale NSD estimates with large-scale coverage. Since all these methods utilize observationally-derived data in parameterization, the emergent/dynamic properties of groundwater can be effectively represented in the framework of LM4-SHARC.

5. Conclusions

817   The presented new framework LM4-SHARC harnesses the parametric efficiency of the DF approximation-

818   based Boussinesq groundwater in capturing the emergent properties of the catchment-scale groundwater. This study

819   proposed a calibration method for the catchment-scale groundwater based on the accuracy of baseflow fluxes and

820   also demonstrated the contribution of the additional groundwater domain/processes (with tuned groundwater

821   parameters) to the improvements of the catchment-scale water/energy budgets. The streamflow recession analysis

822   provides a physically explicit and viable way to use readily available streamflow measurements to infer the time-

823   evolving groundwater properties. Thus, we argued how the time-evolving groundwater diffusivity could be

824   considered in Earth system modeling through the combined use of numerical/explicit groundwater domain and the

825   observationally-derived stream discharge information.

826   The notable improvement in soil moisture and temperature predictions resulting from the LM4-SHARC's

827   hydraulic continuum scheme underscores the necessity of resolving the catchment-scale groundwater dynamics and

828   its interactions with the soil and the stream at the grid-scale of ESM. Specifically, our analysis shows that vertical

829   soil drainage to relatively deep groundwater should be taken into account when simulating soil moisture and

830   temperature. We also note that the soil columns in ESMs might hold too much water without the groundwater-

831   induced drainage dynamics. The significant amounts of facilitated drainage (i.e., ΔSWS) roughly around 90-110 mm

832   y^{-1}, which corresponds to about 6-7 % of the total precipitation in our study site, also emphasizes the importance of

833   considering the lateral groundwater divergence (from hilltop) and convergence (to riparian zone) in ESM land

834   components. The existing biases in soil moisture and surface temperature can lead to a flawed description of other

835   land variables, impacting the surface energy balance, carbon cycle, and biogeochemistry. As the simulated soil

836   temperature was found to be lower than observed values due to the ratio of sensible heat to latent heat and soil heat

837   capacity as a function of SMC, liquid water contained in the partially saturated soil needs to be better quantified as it

838   significantly influences the dynamics of land surface energy balance.

839   Scaling the fine-scale surface water-energy processes to the GCMs' grid cell scales while properly

840   considering the hydrologic interactions and heterogeneity is one of the primary objectives in the ESM community.

841   Considering that the streamflow is a major water flux (that significantly affects energy, carbon, and biogeochemical

842   fluxes) crossing the catchments, in order to properly scale the effects of the SHARC's catchment-scale hydraulic

843   continuum to the macroscale grid cell, a reach-to-reach connection throughout the river network (i.e., stream/river

844   routing) needs to be established. Also, based on the enhanced baseflow production in LM4-SHARC, we expect

845   significant qualitative enhancements of streamflow estimates, which will, in turn, lead to enhanced surface/near-

846   surface water and energy budgets as well as flooding representation (e.g., floodplain dynamics). Overall, the

847   improved water and energy balance in LM4-SHARC is expected to be relevant for coupled land-atmosphere

848   simulations, where refining the land surface state plays a significant role in developing the lower atmospheric

849   boundary layer, and also contributes to the efforts to address societal challenges by hydrologic extreme events such

850   as flooding with enhanced streamflow production.

851

852   **Appendix A: The dynamic heat capacity of the groundwater and vertical heat conduction ($\delta_{v_{cnd}}$) and**

853   **advection ($\delta_{v_{adv}}$) between the soil column and groundwater**

854   LM4-SHARC simulates the heat capacity of the groundwater (hc_{gw}) dynamically according to liquid (i.e.,

855   soil moisture) (wl_{gw}) and frozen water content (i.e., soil ice) (ws_{gw}) in the groundwater. LM4-SHARC represents

856   two mechanisms of heat transfer, advection ($\delta_{v_{adv}}$) and conduction ($\delta_{v_{cnd}}$), between the soil column and

857   groundwater based on the dynamically estimated groundwater heat capacity.

858    $$\delta_{v_{cnd,i}}^{j} = \lambda_{stog,i}^{j}\left(T_{gw,i}^{j} - T_{btm,i}^{j}\right) \tag{A1}$$

859    $$\delta_{v_{adv,i}}^{j} = \begin{cases} \frac{r_i^j}{\rho}\left(T_{btm,i}^{j} - T_{freeze}\right)\frac{hc_{gw,i}^{j}}{\Delta t \Delta L_i^j} & (r > 0) \\[2ex] \frac{r_i^j}{\rho}\left(T_{gw,i}^{j} - T_{freeze}\right)\frac{hc_{gw,i}^{j}}{\Delta t \Delta L_i^j} & (r \leq 0) \end{cases} \tag{A2}$$

860   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

861   groundwater (J m^{-2} s^{-1}). $\lambda_{stog,i}^{j}$ is the thermal transmittance (W m^{-2} K^{-1}) between the bottommost soil layer and

862   groundwater at i th height band at j th time step. $T_{gw,i}^{j} - T_{btm,i}^{j}$ is the temperature difference between the bottommost

863   layer and groundwater.

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

865   where wl_{gw}, ws_{gw} are the liquid and ice content in the groundwater, respectively (-). $hc_{btm,i}^{j}$ the dry soil heat

866   capacity of the bottommost soil layer at i th height band at j th time step (J K^{-1} m^{-2}). clw is the specific heat of liquid

867   (4218.0 J kg^{-1} K^{-1}), and csw is the specific heat of ice (2106.0 J kg^{-1} K^{-1}). We note that each HB-scale groundwater

868   domain's thermal properties were assumed to be identical to those of the corresponding soil column's bottommost

869   soil layer.

870

871   ***Author contributions***

872   All authors contributed to research design and manuscript editing. M.H. developed software, performed the

873   model evaluation, data analysis and drafted the first version of this manuscript. N.C., S.M., and E.S. contributed to

874   software development. E.Z., and A.P. contributed to data preparation.

875

876   ***Competing interests***

877   A co-author is a member of the editorial board of Geoscientific Model Development (GMD).

878

879   ***Acknowledgement***

880   This report was prepared by Minki Hong under award NA23OAR4050431I from the National Oceanic and

881   Atmospheric Administration, U.S. Department of Commerce. The statements, findings, conclusions, and

882   recommendations are those of the author(s) and do not necessarily reflect the views of the National Oceanic and

883 Atmospheric Administration, or the U.S. Department of Commerce. The authors thank Dr. John Dunne and Dr.

884 Kirsten Findell at NOAA GFDL for reviewing a first draft of the manuscript.

885

886 *Code and data availability*

887   The source code of LM4-SHARC v1.0 and the model input data such as model domain dataset and forcing

888 data are shared in a public repository: https://zenodo.org/records/13750071.

889

890 **References**

891 Andreae, M. O., Artaxo, P., Brandao, C., Carswell, F. E., Ciccioli, P., Da Costa, A., Culf, A. D., Esteves, J., Gash, J.
892   H., and Grace, J.: Biogeochemical cycling of carbon, water, energy, trace gases, and aerosols in Amazonia:
893   The LBA-EUSTACH experiments, J. Geophys. Res.: Atmos., 107, LBA 33-31-LBA 33-25, 2002.
894 Bales, R., Stacy, E., Safeeq, M., Meng, X., Meadows, M., Oroza, C., Conklin, M., Glaser, S., and Wagenbrenner, J.:
895   Spatially distributed water-balance and meteorological data from the rain–snow transition, southern Sierra
896   Nevada, California, Earth Syst. Sci. Data, 10, 1795-1805, 10.5194/essd-10-1795-2018, 2018.
897 Baratelli, F., Flipo, N., Rivière, A., and Biancamaria, S.: Retrieving river baseflow from SWOT spaceborne mission,
898   Remote Sens. Environ., 218, 44-54, 2018.
899 Baroni, G., Schalge, B., Rakovec, O., Kumar, R., Schüler, L., Samaniego, L., Simmer, C., and Attinger, S.: A
900   Comprehensive Distributed Hydrological Modeling Intercomparison to Support Process Representation and
901   Data Collection Strategies, Water Resour. Res., 55, 990-1010, 10.1029/2018wr023941, 2019.
902 Bart, R., and Hope, A. (2014). Inter-seasonal variability in baseflow recession rates: The role of aquifer antecedent
903   storage in central California watersheds. J. Hydrol., 519, 205-213.
904 Basha, H. A.: Traveling wave solution of the Boussinesq equation for groundwater flow in horizontal aquifers,
905   Water Resour. Res., 49, 1668-1679, 10.1002/wrcr.20168, 2013.
906 Best, M. J., Pryor, M., Clark, D., Rooney, G. G., Essery, R., Ménard, C., Edwards, J., Hendry, M., Porson, A., and
907   Gedney, N.: The Joint UK Land Environment Simulator (JULES), model description–Part 1: energy and
908   water fluxes, Geosci. Model Dev., 4, 677-699, 2011.
909 Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology/Un modèle
910   à base physique de zone d'appel variable de l'hydrologie du bassin versant, Hydrol. Sci. J., 24, 43-69, 1979.
911 Bisht, G., Huang, M. Y., Zhou, T., Chen, X. Y., Dai, H., Hammond, G. E., Riley, W. J., Downs, J. L., Liu, Y., and
912   Zachara, J. M.: Coupling a three-dimensional subsurface flow and transport model with a land surface
913   model to simulate stream-aquifer-land interactions (CP v1.0), Geosci. Model Dev., 10, 4539-4562,
914   10.5194/gmd-10-4539-2017, 2017.
915 Blyth, E. M., Arora, V. K., Clark, D. B., Dadson, S. J., De Kauwe, M. G., Lawrence, D. M., Melton, J. R., Pongratz,
916   J., Turton, R. H., Yoshimura, K., and Yuan, H.: Advances in Land Surface Modelling, Curr. Clim. Change
917   Rep., 7, 45-71, 10.1007/s40641-021-00171-5, 2021.
918 Boussinesq, J.: Sur le de´bit, en temps de se´cheresse, d'une source alimente´e par une nappe d'eaux d'infiltration,
919   C.R. Hebd Se´-ances Acad. Sci., 136, 1511-1517, 1903.
920 Brutsaert, W.: The unit response of groundwater outflow from a hillslope, Water Resour. Res., 30, 2759-2763,
921   10.1029/94wr01396, 1994.
922 Brutsaert, W. and Lopez, J. P.: Basin-scale geohydrologic drought flow features of riparian aquifers in the Southern
923   Great Plains, Water Resour. Res., 34, 233-240, 10.1029/97wr03068, 1998.
924 Brutsaert, W. and Nieber, J. L.: Regionalized drought flow hydrographs from a mature glaciated plateau, Water
925   Resour. Res., 13, 637-643, 10.1029/wr013i003p00637, 1977.
926 Cai, H., Shi, H., Liu, S., and Babovic, V.: Impacts of regional characteristics on improving the accuracy of
927   groundwater level prediction using machine learning: The case of central eastern continental United States,
928   J. Hydrol.: Reg. Stud., 37, 100930, 2021.
929 Campoy, A., Ducharne, A., Cheruy, F., Hourdin, F., Polcher, J., and Dupont, J.: Response of land surface fluxes and
930   precipitation to different soil bottom hydrological conditions in a general circulation model, J. Geophys.
931   Res.: Atmos., 118, 10,725-710,739, 2013.
932 Chaney, N. W., Metcalfe, P., and Wood, E. F.: HydroBlocks: a field-scale resolving land surface model for
933   application over continental extents, Hydrol. Processes, 30, 3543-3559, 10.1002/hyp.10891, 2016.

934   Chaney, N. W., Torres-Rojas, L., Vergopolan, N., and Fisher, C. K.: HydroBlocks v0.2: enabling a field-scale two-
935   way coupling between the land surface and river networks in Earth system models, Geosci. Model Dev.,
936   14, 6813-6832, 10.5194/gmd-14-6813-2021, 2021.
937   Chaney, N. W., Van Huijgevoort, M. H. J., Shevliakova, E., Malyshev, S., Milly, P. C. D., Gauthier, P. P. G., and
938   Sulman, B. N.: Harnessing big data to rethink land heterogeneity in Earth system models, Hydrol. Earth
939   Syst. Sci., 22, 3311-3330, 10.5194/hess-22-3311-2018, 2018.
940   Chen, X. and Hu, Q.: Groundwater influences on soil moisture and surface evaporation, J. Hydrol., 297, 285-300,
941   2004.
942   Cheng, S., Cheng, L., Liu, P., Qin, S., Zhang, L., Xu, C. Y., Xiong, L., Liu, L., and Xia, J.: An Analytical Baseflow
943   Coefficient Curve for Depicting the Spatial Variability of Mean Annual Catchment Baseflow, Water
944   Resour. Res., 57, 10.1029/2020wr029529, 2021.
945   Clark, M. P., Rupp, D. E., Woods, R. A., Tromp-Van Meerveld, H. J., Peters, N. E., and Freer, J. E.: Consistency
946   between hydrological models and field observations: linking processes at the hillslope scale to hydrological
947   responses at the watershed scale, Hydrol. Processes, 23, 311-319, 10.1002/hyp.7154, 2009.
948   Clark, M. P., Fan, Y., Lawrence, D. M., Adam, J. C., Bolster, D., Gochis, D. J., Hooper, R. P., Kumar, M., Leung,
949   L. R., Mackay, D. S., Maxwell, R. M., Shen, C. P., Swenson, S. C., and Zeng, X. B.: Improving the
950   representation of hydrologic processes in Earth System Models, Water Resour. Res., 51, 5929-5956,
951   10.1002/2015wr017096, 2015.
952   De Rosnay, P., Polcher, J. d., Bruen, M., and Laval, K.: Impact of a physically based soil water flow and soil-plant
953   interaction representation for modeling large-scale land surface processes, J. Geophys. Res.: Atmos., 107,
954   ACL 3-1-ACL 3-19, 2002.
955   De St Venant, B.: Theorie du mouvement non-permanent des eaux avec application aux crues des rivers et a
956   l'introdunion des Marees dans leur lit, Academic de Sci. Comptes Redus, 73, 148-154, 1871.
957   Delworth, T. L., Cooke, W. F., Adcroft, A., Bushuk, M., Chen, J. H., Dunne, K. A., Ginoux, P., Gudgel, R.,
958   Hallberg, R. W., and Harris, L.: SPEAR: The next generation GFDL modeling system for seasonal to
959   multidecadal prediction and projection, J. Adv. Model. Earth Syst., 12, e2019MS001895, 2020.
960   Dunne, J. P., Horowitz, L., Adcroft, A., Ginoux, P., Held, I., John, J., Krasting, J. P., Malyshev, S., Naik, V., and
961   Paulot, F.: The GFDL Earth System Model version 4.1 (GFDL-ESM 4.1): Overall coupled model
962   description and simulation characteristics, J. Adv. Model. Earth Syst., 12, e2019MS002015, 2020.
963   Dupuit, J.: Etudes theoriques et pratiques sur le mouvement des eaux dans les canaux decouverts et a travers les
964   terrains permeables avec des considerations relatives au regime des grandes eaux, au debouche a leur
965   donner et a la marche des alluvions dans les rivieres a fond mobile, Dunod1863.
966   Fan, Y.: Groundwater in the Earth's critical zone: Relevance to large-scale patterns and processes, Water Resour.
967   Res., 51, 3052-3069, 10.1002/2015wr017037, 2015.
968   Fan, Y. and Miguez-Macho, G.: A simple hydrologic framework for simulating wetlands in climate and earth
969   system models, Clim. Dyn., 37, 253-278, 2011.
970   Fan, Y., Miguez-Macho, G., Weaver, C. P., Walko, R., and Robock, A.: Incorporating water table dynamics in
971   climate modeling: 1. Water table observations and equilibrium water table simulations, Journal of
972   Geophysical Research-Atmospheres, 112, 10.1029/2006jd008111, 2007.
973   Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., Brooks, P. D., Dietrich, W. E.,
974   Flores, A., Grant, G., Kirchner, J. W., Mackay, D. S., McDonnell, J. J., Milly, P. C. D., Sullivan, P. L.,
975   Tague, C., Ajami, H., Chaney, N., Hartmann, A., Hazenberg, P., McNamara, J., Pelletier, J., Perket, J.,
976   Rouholahnejad-Freund, E., Wagener, T., Zeng, X., Beighley, E., Buzan, J., Huang, M., Livneh, B.,
977   Mohanty, B. P., Nijssen, B., Safeeq, M., Shen, C., Verseveld, W., Volk, J., and Yamazaki, D.: Hillslope
978   Hydrology in Global Change Research and Earth System Modeling, Water Resour. Res., 55, 1737-1772,
979   10.1029/2018wr023903, 2019.
980   Forchheimer, P.: Uber die ergiebigkeit von brummen, Anlagen und Sickerschlitzen, Zeitsch Archit. Ing. Ver.,
981   Hannover, 32, 539-563, 1986.
982   Gentine, P., Green, J. K., Guérin, M., Humphrey, V., Seneviratne, S. I., Zhang, Y., and Zhou, S.: Coupling between
983   the terrestrial carbon and water cycles—a review, Environ. Res. Lett., 14, 083003, 2019.
984   Gleeson, T., Wang-Erlandsson, L., Porkka, M., Zipper, S. C., Jaramillo, F., Gerten, D., Fetzer, I., Cornell, S. E.,
985   Piemontese, L., Gordon, L. J., Rockström, J., Oki, T., Sivapalan, M., Wada, Y., Brauman, K. A., Flörke,
986   M., Bierkens, M. F. P., Lehner, B., Keys, P., Kummu, M., Wagener, T., Dadson, S., Troy, T. J., Steffen,
987   W., Falkenmark, M., and Famiglietti, J. S.: Illuminating water cycle modifications and Earth system
988   resilience in the Anthropocene, Water Resour. Res., 56, 10.1029/2019wr024957, 2020.

Gochis, D. J., M. Barlage, A. Dugger, K. FitzGerald, L. Karsten, M. McAllister, J. McCreight, J. Mills, A. RafieeiNasab, L. Read, K. Sampson, D. Yates, W. Yu,: The NCAR WRF-Hydro Modeling System Technical Description (Version 5.2), National Center for Atmospheric Research, 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 Survey2004.

Hong, M., Mohanty, B.P., Sheng, Z., 2020. An Explicit Scheme to Represent the Bidirectional Hydrologic Exchanges Between the Vadose Zone, Phreatic Aquifer, and River. Water Resour. Res., 56.. https://doi.org/10.1029/2020wr027571

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, 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 Syst., 15, 10.1029/2022ms003325, 2023b.

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-+, 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 Nevada1, JAWRA J. Am. Water Resour. Assoc., 48, 667-678, 10.1111/j.1752-1688.2012.00641.x, 2012.

Jachens, E. R., Roques, C., Rupp, D. E., and Selker, J. S.: Recession analysis revisited: impacts of climate on parameter estimation, Hydrol. Earth Syst. Sci., 24, 1159-1170, 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, 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, n/a-n/a, 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, 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, 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, Scientific Data, 10, 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 Syst., 3, 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., Lombardozzi, 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 Syst., 11, 4245-4287, 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. Dyn., 36, 57-81, 2011.

1044   Li, D., Andreadis, K. M., Margulis, S. A., and Lettenmaier, D. P.: A Data Assimilation Framework for Generating
1045   Space-Time Continuous Daily SWOT River Discharge Data Products, Water Resour. Res., 56,
1046   10.1029/2019wr026999, 2020.
1047   Li, H., and Ameli, A. (2022). A statistical approach for identifying factors governing streamflow recession
1048   behaviour. Hydrol. Process., 36(10), e14718.
1049   Li, H., Wigmosta, M. S., Wu, H., Huang, M., Ke, Y., Coleman, A. M., and Leung, L. R.: A physically based runoff
1050   routing model for land surface and earth system models, J. Hydrometeorol., 14, 808-828, 2013.
1051   Liang, X., Xie, Z. H., and Huang, M. Y.: A new parameterization for surface and groundwater interactions and its
1052   impact on water budgets with the variable infiltration capacity (VIC) land surface model, J. Geophys. Res.:
1053   Atmos., 108, 10.1029/2002jd003090, 2003.
1054   Manning, R., Griffith, J. P., Pigot, T., and Vernon-Harcourt, L. F.: On the flow of water in open channels and pipes,
1055   1890.
1056   Maxwell, R. M. and Kollet, S. J.: Interdependence of groundwater dynamics and land-energy feedbacks under
1057   climate change, Nat. Geosci., 1, 665-669, 10.1038/ngeo315, 2008.
1058   Maxwell, R. M., Chow, F. K., and Kollet, S. J.: The groundwater–land-surface–atmosphere connection: Soil
1059   moisture effects on the atmospheric boundary layer in fully-coupled simulations, Adv. Water Resour., 30,
1060   2447-2466, 10.1016/j.advwatres.2007.05.018, 2007.
1061   Maxwell, R. M., Condon, L. E., and Kollet, S. J.: A high-resolution simulation of groundwater and surface water
1062   over most of the continental US with the integrated hydrologic model ParFlow v3, Geosci. Model Dev., 8,
1063   923-937, 10.5194/gmd-8-923-2015, 2015.
1064   Maxwell, R. M., Lundquist, J. K., Mirocha, J. D., Smith, S. G., Woodward, C. S., and Tompson, A. F. B.:
1065   Development of a Coupled Groundwater–Atmosphere Model, Mon. Weather Rev., 139, 96-116,
1066   10.1175/2010mwr3392.1, 2011.
1067   McCabe, M., Wood, E. F., Wójcik, R., Pan, M., Sheffield, J., Gao, H., and Su, H.: Hydrological consistency using
1068   multi-sensor remote sensing data for water and energy cycle studies, Remote Sens. Environ., 112, 430-444,
1069   2008.
1070   Miguez-Macho, G., Fan, Y., Weaver, C. P., Walko, R., and Robock, A.: Incorporating water table dynamics in
1071   climate modeling: 2. Formulation, validation, and soil moisture simulation, J. Geophys. Res.: Atmos., 112,
1072   10.1029/2006jd008112, 2007.
1073   Miguez-Macho, G. and Fan, Y.: The role of groundwater in the Amazon water cycle: 1. Influence on seasonal
1074   streamflow, flooding and wetlands, J. Geophys. Res.: Atmos., 117, n/a-n/a, 10.1029/2012jd017539, 2012.
1075   Milly, P. C., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., and Stouffer, R.
1076   J.: Stationarity is dead: Whither water management?, Science, 319, 573-574, 2008.
1077   Milly, P. C. D., Malyshev, S. L., Shevliakova, E., Dunne, K. A., Findell, K. L., Gleeson, T., Liang, Z., Phillipps, P.,
1078   Stouffer, R. J., and Swenson, S.: An Enhanced Model of Land Water and Energy for Global Hydrologic
1079   and Earth-System Studies, J. Hydrometeorol., 15, 1739-1761, 10.1175/jhm-d-13-0162.1, 2014.
1080   Mu, Q., Zhao, M., and Running, S. W.: Evolution of hydrological and carbon cycles under a changing climate,
1081   Hydrol. Processes, 25, 4093-4102, 10.1002/hyp.8367, 2011.
1082   Newman, A. J., Clark, M. P., Winstral, A., Marks, D., and Seyfried, M.: The Use of Similarity Concepts to
1083   Represent Subgrid Variability in Land Surface Models: Case Study in a Snowmelt-Dominated Watershed,
1084   J. Hydrometeorol., 15, 1717-1738, 10.1175/jhm-d-13-038.1, 2014.
1085   Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D.,
1086   Rosero, E., Tewari, M., and Xia, Y.: The community Noah land surface model with multiparameterization
1087   options (Noah-MP): 1. Model description and evaluation with local-scale measurements, J. Geophys. Res.,
1088   116, 10.1029/2010jd015139, 2011.
1089   Oleson, K. W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., Levis, S., Li, F., Riley, W.
1090   J., Subin, Z. M., Swenson, S. C., and Thornton, P. E.: Technical Description of version 4.5 of the
1091   Community Land Model (CLM), NCAR Earth System Laboratory, 10.5065/D6RR1W7M, 2013.
1092   Ouyang, L., Liu, S., Ye, J., Liu, Z., Sheng, F., Wang, R., and Lu, Z.: Quantitative assessment of surface runoff and
1093   base flow response to multiple factors in Pengchongjian small watershed, Forests, 9, 553, 2018.
1094   Poggio, L., De Sousa, L.M., Batjes, N.H., Heuvelink, G.B.M., Kempen, B., Ribeiro, E., and Rossiter, D., 2021.
1095   SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. SOIL 7, 217–
1096   240.. https://doi.org/10.5194/soil-7-217-2021
1097   Pappenberger, F., Dutra, E., Wetterhall, F., and Cloke, H. L.: Deriving global flood hazard maps of fluvial floods
1098   through a physical model cascade, Hydrol. Earth Syst. Sci., 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, 10.1002/2015wr017749, 2015.

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, 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, n/a-n/a, 10.1029/2006wr005080, 2006.

Safeeq, M. and Hunsaker, C. T.: Characterizing Runoff and Water Yield for Headwater Catchments in the Southern Sierra Nevada, JAWRA J. Am. Water Resour. Assoc., 52, 1327-1346, 10.1111/1752-1688.12457, 2016.

Shangguan, W., Dai, Y., Duan, Q., Liu, B., and Yuan, H., 2014. A global soil data set for earth system modeling. J. Adv. Model. Earth Syst., 6, 249–263.. https://doi.org/10.1002/2013ms000293

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, 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., Phillipps, 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 Syst., 16, 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, 2006.

Strack, O. D. L.: A Dupuit-Forchheimer Model for three-dimensional flow with variable density, Water Resour. Res., 31, 3007-3017, 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, 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, Scientific Data, 9, 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, 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, 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, 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 Syst., 13, 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, 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, 10.1002/wrcr.20407, 2013.

Trotter, L., Saft, M., Peel, M. C., and Fowler, K. J. (2024). Recession constants are non-stationary: Impacts of multi-annual drought on catchment recession behaviour and storage dynamics. J. Hydrol., 630, 130707.

USGS: 3D Elevation Program 1-Meter Resolution Digital Elevation Model [dataset], 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. Processes, 28, 6276-6291, 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, 2020.

1154 Xu, B., Ye, M., Dong, S., Dai, Z., and Pei, Y.: A new model for simulating spring discharge recession and
1155   estimating effective porosity of karst aquifers, J. Hydrol., 562, 609-622, 2018.
1156 Xu, Z., Siirila-Woodburn, E. R., Rhoades, A. M., and Feldman, D.: Sensitivities of subgrid-scale physics schemes,
1157   meteorological forcing, and topographic radiation in atmosphere-through-bedrock integrated process
1158   models: a case study in the Upper Colorado River basin, Hydrol. Earth Syst. Sci., 27, 1771-1789, 2023.
1159 Yeh, P. J. F. and Eltahir, E. A. B.: Representation of Water Table Dynamics in a Land Surface Scheme. Part I:
1160   Model Development, Journal of Climate, 18, 1861-1880, 10.1175/jcli3330.1, 2005.
1161 Zecharias, Y. B. and Brutsaert, W.: Recession characteristics of groundwater outflow and base flow from
1162   mountainous watersheds, Water Resour. Res., 24, 1651-1658, 1988.
1163 Zeng, Y. J., Xie, Z. H., Yu, Y., Liu, S., Wang, L. Y., Jia, B. H., Qin, P. H., and Chen, Y. N.: Ecohydrological effects
1164   of stream-aquifer water interaction: a case study of the Heihe River basin, northwestern China, Hydrol.
1165   Earth Syst. Sci., 20, 2333-2352, 10.5194/hess-20-2333-2016, 2016.
1166 Zhao, M., Golaz, J. C., Held, I., Guo, H., Balaji, V., Benson, R., Chen, J. H., Chen, X., Donner, L., and Dunne, J.:
1167   The GFDL global atmosphere and land model AM4. 0/LM4. 0: 2. Model description, sensitivity studies,
1168   and tuning strategies, J. Adv. Model. Earth Syst., 10, 735-769, 2018.
1169 Zorzetto, E., Malyshev, S., Chaney, N., Paynter, D., Menzel, R., and Shevliakova, E.: Effects of complex terrain on
1170   the shortwave radiative balance: a sub-grid-scale parameterization for the GFDL Earth System Model
1171   version 4.1, Geosci. Model Dev., 16, 1937-1960, 10.5194/gmd-16-1937-2023, 2023.