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

37   water cycle, the groundwater (i.e., water in saturated zones beneath the land surface) plays a pivotal role in

38   developing the surface thermal energy and moisture dynamics and the land-atmosphere coupling by regulating how

39   the water and thermal energy is stored and transported across the landscape (Andreae et al., 2002; Mu et al., 2011;

40   Mccabe et al., 2008; Gentine et al., 2019; Fan, 2015). The effects of the groundwater state (e.g., storage) on the

41   water-energy balance at the land surface have been discussed in several studies based on an explicitly treated

42   groundwater scheme (Miguez-Macho et al., 2007; Liang et al., 2003; Yeh and Eltahir, 2005; Zeng et al., 2016;

43   Maxwell et al., 2011; Maxwell and Kollet, 2008). The studies commonly found higher sensitivity of surface energy

44   balance to groundwater storage if the water table is shallow (e.g., the water table depth less than 5 m).

45   However, many Earth system models (ESMs) only represent a few top meters of soil, and most, if not all,

46   ignore two-way interaction between the groundwater and other components of the hydrologic cycle. Specifically,

47   ESMs lack proper land surface scheme(s) to represent the two-way interactions between the groundwater and

48   stream/river while conserving the hydraulic continuity between them. In most cases, the river routing module

49   implemented in the ESMs exists mainly for their traditional purposes (e.g., linking precipitation-induced runoff to

50   the ocean) without the ability to consider the groundwater variations driven by the horizontal hydraulic gradients

51   between the stream and groundwater (Li et al., 2013; De Rosnay et al., 2002; Lawrence et al., 2011; Lawrence et al.,

52   2019; Miguez-Macho and Fan, 2012; Best et al., 2011; Takata et al., 2003).

53   Furthermore, the resolution of climate/water-related information provided by the current ESMs is too low

54   to characterize hydrological extremes and address the stakeholders' need to evaluate potential impact of future

55   climate change. Even an atmosphere-land water information produced at the resolution of 0.25° (Delworth et al.,

56   2020) — considered the most high-resolution in ESMs — is still not sufficiently fine to generate fine-scale data for

57   decision-making. The land components of ESMs generate the streamflow estimates only at a grid-scale (typically

58   ranging from 0.5° to 1.0°) (Oleson et al., 2013; Miguez-Macho et al., 2007; Fan et al., 2007; Pappenberger et al.,

59   2012; Campoy et al., 2013) and since such coarse-resolution streamflow data, for example, is not suitable to directly

60   use for locating hydrologic events at a fine-scale even in case of extremes (e.g., flooding), the scale mismatch

61   (between the stakeholders and ESMs) could reduce the usability of the ESM-derived projections. Another challenge

62   lies in the high degree of land heterogeneity with respect to soil and topographic characteristics, which needs to be

63   resolved at the ESMs' grid-scale. Considering the significant impacts of the fine-scale soil/topographic properties on

64   the hydrologic processes, the two-way interactions between the groundwater and the rest of the hydrologic cycle

65   must be parameterized at the sub-grid scale (e.g., catchment) to adequately represent the sub-grid variability of

66   hydrologic states and its implication for the broader Earth system processes and interactions (Gleeson et al., 2020;

67   Xu et al., 2023; Maxwell and Kollet, 2008; Pokhrel et al., 2013).

68   Catchments provide an appropriate scale to properly capture such sub-grid spatial heterogeneity and its

69   effect on the interactions among the hydrological cycle components (Clark et al., 2015; Fan et al., 2019; Blyth et al.,

70   2021). This is mainly due to the suitability of the catchment-scale in describing the topographically-driven flow

71   characteristics of water and other major fluxes (e.g., thermal energy) caused by water transport. In fact, catchments

72   have been considered hydrologic spatial units (or response units) where the theoretical conceptualization of surface

73   and subsurface water transport can be tested with readily available observational data, such as streamflow

74   measurements (Sivapalan, 2006; Troch et al., 2013; Kirchner, 2009). Over the past decades, many efforts have been

75   made to depict the hydraulically-based interactions among the distinct flow domains at the hillslope and catchment

76   scales (Kollet and Maxwell, 2006; Niu et al., 2011; Shen and Phanikumar, 2010; Gochis, 2021). These efforts

77   mainly aimed at representing the water-energy coupled balance while considering the time-dependent non-linear

78   relationship between water states and fluxes based on Darcian flow (Clark et al., 2015; Fan et al., 2019).

79   However, more specifically concerning the representation of the groundwater and its interactions with the

80   overlying soil and the stream/river, the spatial heterogeneity of groundwater properties (e.g., diffusivity, or effective

81   (drainable) porosity) remains a significant challenge for enhancing the predictability of terrestrial water storage and

82   exchange fluxes (Clark et al., 2009; Jing et al., 2019). Notably, an accurate description of groundwater

83   properties/processes at the catchment-scale is crucial because the hydrologic convergence (to the riparian zone) and

84   divergence (from the hilltops) by the groundwater movements significantly affect the catchments' water and energy

85   dynamics (Fan, 2015; Miguez-Macho and Fan, 2012; Chen and Hu, 2004; Maxwell et al., 2007). Even in existing

86   studies to capture the interactions among the soil-groundwater-stream, the properties of the groundwater were

87   treated as constant and not as emergent and time-evolving properties affected by climate or human activities (e.g.,

88   land use and cover changes). For example, these studies empirically fit coefficients for a storage-discharge function

89   (e.g., bucket model) or predefined soil properties dataset to parameterize the groundwater domains (Zeng et al.,

90   2016; Gochis, 2021; Newman et al., 2014; Kollet and Maxwell, 2008; Leung et al., 2011). Thus, a theoretical

91   approach to capture the dynamic, emergent properties of the catchment-scale groundwater has been largely absent in

92   those previous efforts.

93   The Dupuit-Forchheimer (DF) approximation allows the incorporation of the catchment-scale groundwater

94   into ESMs and accounts for its emergent properties/processes shaping the interactions with the unsaturated soil and

95   the stream (Dupuit, 1863; Forchheimer, 1986). According to the DF approximation, heterogeneity in groundwater

96   properties can be represented by effective parameters reflecting combined effects of the groundwater properties

97   variability on fluxes (e.g., baseflow), assuming that the lateral groundwater discharge is from the

98   homogeneous/isotropic (or systemically-defined) groundwater (Rupp and Selker, 2006; Strack, 1995; Beven and

99   Kirkby, 1979). The DF approximation is especially relevant as the heterogeneity of groundwater properties not

100   representable below model-specific resolution can be considered one of the challenges to the accuracy of modeled

101   data (Baroni et al., 2019; Maxwell et al., 2015; Niu et al., 2011; Bisht et al., 2017). The DF approximation is derived

102   by recasting the analytical solution of the Boussinesq equation (Boussinesq, 1903) into the power-law streamflow

103   recession model by Brutsaert and Nieber (1977). The method of streamflow recession analysis provides a theoretical

104   basis for estimating the effective groundwater properties on a catchment scale using (relatively readily) available

105   streamflow observations (Tashie et al., 2021; Hong and Mohanty, 2023a; Vannier et al., 2014; Brutsaert and Lopez,

106   1998; Troch et al., 1993; Xu et al., 2018).

107   The catchment-scale processes can be resolved at the ESMs' grid-scale via approaches to the sub-grid

108   model structure such as HydroBlocks (Chaney et al., 2016; Chaney et al., 2018). The hierarchical multivariate

109   clustering (HMC) approach, as implemented in LM4-HydroBlocks (Chaney et al., 2018), enables partitioning a

110   macroscale land domain grid (e.g., 0.5°, 1.0°) into sub-grid terrain units (termed 'tiles') (Milly et al., 2014; Zhao et

111   al., 2018; Dunne et al., 2020). Then, the catchment-scale hydrologic structure of the macroscale grid cell is

112   described based on the inter-tile connection enforced through the flow accumulation area derived from the digital

113   elevation model (DEM) (Chaney et al., 2018). This study presents a new parameterization of two-way Soil-Hillslope

114   Aquifer-River Continuum, namely SHARC (v1.0), tailored to the fourth generation of the Geophysical Fluid

115   Dynamics Laboratory (GFDL) land model version 4 (i.e., LM4) aiming to represent the catchment-scale two-way

116   interactions among the unsaturated soil (i.e., vadose zone), the groundwater, and stream/river in the GFDL Earth

117   system modeling framework. The new LM4-SHARC has been developed by extending the existing hillslope

118   hydrology scheme currently used in LM4-HydroBlocks (Subin et al., 2014; Chaney et al., 2018), and relying on the

119   catchment-scale hydrologic partitioning (of the macroscale grid cell) by HMC approach for the explicit treatment of

120   divergence fluxes among soil-groundwater-stream.

121   In sum, the LM4-SHARC (1) explicitly characterizes the catchment-scale groundwater while accounting

122   for its emergent properties such as groundwater diffusivity, (2) represents the two-way water and energy exchanges

123   in the vertical direction between the soil–groundwater and in the lateral direction between the groundwater–stream,

124   and (3) accounts for the groundwater-induced variations in surface water-energy budgets to enhance the ESM's

125   realism of land hydrology. This study primarily compares the two model configurations to evaluate the importance

126   of accurately describing the catchment-scale groundwater in the land component of ESM. As a proof of concept, we

127   apply the LM4-SHARC and LM4-HydroBlocks to a headwater catchment in the Sierra National Forest, Nevada, and

128   evaluate the modeled outputs against corresponding observations. Additionally, we discuss how the hydrologic

129   convergence to the river valley (and divergence from the hilltop) of liquid water and thermal energy due to the

130   groundwater flow could contribute to hydrologic contrast in a catchment based on simulations.

131

132   **2.   Methods**

133   **2.1.   Representation of the soil-groundwater processes in LM4-HydroBlocks**

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

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

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

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

138   derived (Subin et al., 2014; Milly et al., 2014). Each soil column in a catchment is composed of a vertical 1-D model

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

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

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

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

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

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

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

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

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

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

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

150   The water fluxes between adjacent tiles are simulated on a layer-to-layer basis, defined following the soil

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

152 layers. The inter-tile heat and dissolved organic carbon fluxes advected by water transport are also represented.

153 While the tile located right next to the river reach interacts with the river, the flux from the tile to the river is a one-

154 way flux as the reach-scale streams' (corresponding to a catchment) hydrograph are not accounted for in LM4-

155 HydroBlocks. With respect to the flux exchange between the river and the soil–bedrock column (adjacent to the

156 reach), in fact, the river stage is set to zero (static pressure head in the reach). Consequently, the water flux is

157 invariably one-way, from the hillslope to the river.

158

159 **2.2. Representation of the Soil-Hillslope Aquifer-River Interactions in LM4-SHARC**

160 **2.2.1. Two-way water transport and conservation in LM4-SHARC**

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

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

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

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

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

166 1986). Therefore, the lateral hydraulic gradient is the only driver of groundwater lateral discharge fluxes as the

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

168 in an unconfined groundwater overlaying an impermeable bedrock of slope θ, the lateral groundwater flux is

169 estimated following equation (1).

170 $$q_l = -K_s \left[\cos\theta \left(\frac{\partial N}{\partial x} \right) + \sin\theta \right] \tag{1}$$

171 where q_l is the speed of lateral groundwater divergence flux (mm s⁻¹). K_s is the saturated lateral hydraulic

172 conductivity (mm s⁻¹). Thus, the groundwater flow rate per unit width of the groundwater is given by $q_l N$, where N

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

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

175 $$f \frac{\partial N}{\partial t} = \cos\theta \frac{\partial}{\partial x} \left(K_s N \frac{\partial N}{\partial x} \right) + \sin\theta \frac{\partial}{\partial x} (K_s N) \tag{2}$$

176 where f is the effective porosity of the groundwater (m³ m⁻³). As implied by the Boussinesq equation, the

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

178 denotes the groundwater hydraulic gradient between adjacent tiles. At each tile, the water table is determined by the

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

180 $$H_i^{j+1} = H_i^j + \left(\frac{r_i^j - q_{l_i}^j}{\rho f}\right)\Delta t \qquad (i = 1,\ldots,n_{HB}) \tag{3}$$

181 where H_i^j is the hydraulic head of the water table (m) at i th height band (HB) (Figure 1) at j th time step. Based on

182 the continuity equation, the $N - H$ relationship can be established as $H = \frac{N}{\cos\theta}$, n_{HB} is the total number of HBs in the

183 catchment. ρ is r_i^j is the liquid water flux between the soil–bedrock column and the water table (mm s^{-1}) at i th

184 height band at j th time step. The physical time step Δt of the model is 30 minutes. $q_{l_i}^j$ denotes the divergence of

185 lateral groundwater flux at i th height band at j th time step and is set to the difference between the lateral divergence

186 from/to HBs immediately above and below. Particularly, in the HB$_1$ (i.e., the nearest height band to the river reach),

187 $q_{l_1}^j$ is determined by the balance of the groundwater discharge from HB$_2$ and the baseflow (if gaining reach) or the

188 channel infiltration (if losing reach). If multiple tiles exist in an HB, r_i is effectively calculated by arithmetically

189 averaging the r values from each tile belonging to the i th height band.

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

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

192 continuity equation with kinematic wave assumption is solved for the river dynamics. For a kinematic wave, the

193 momentum equation assumes that the energy grade line is parallel to the streambed (Equation 4) (De St Venant,

194 1871; Strelkoff, 1970). Based on Manning's equation, the momentum equation can be recast in the relationship

195 between the steam discharge Q (m^3 s^{-1}) and the cross-section area of flow U (m^2) (Equation 5) (Manning et al.,

196 1890). Together with the continuity equation, the time derivative of the steam discharge ($\frac{dQ}{dt}$) with lateral influx q_l (as

197 defined in Equation 1) can be expressed as Equation 6.

198 $$S_0 = S_f \tag{4}$$

199 $$U = \alpha Q^\beta \tag{5}$$

200 $$\frac{dQ}{dy} = \frac{\partial Q}{\partial y} + \frac{\partial Q}{\partial t}\frac{1}{c_k} = \frac{2q_l H_1}{\rho} \tag{6}$$

201 where c_k is the kinematic wave celerity (i.e., $c_{k = \frac{dQ}{dU}}$ (m s^{-1})), and y denotes the river flow direction coordinate. H_1 is

202 the vertical thickness of the groundwater at HB$_1$, which effectively defines the wetted perimeter at the stream-

203 hillslope interface (m). Then, the reach outflow (i.e., discharge at the catchment outlet) was used to inversely

204 estimate the river stage, which was in turn used to determine the lateral hydraulic gradients between the river stage

205 and the water table in HB$_1$.

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

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

208 domain. As shown in Figure 1, due to the hydraulic gradient between the bottommost soil layer and the water table,
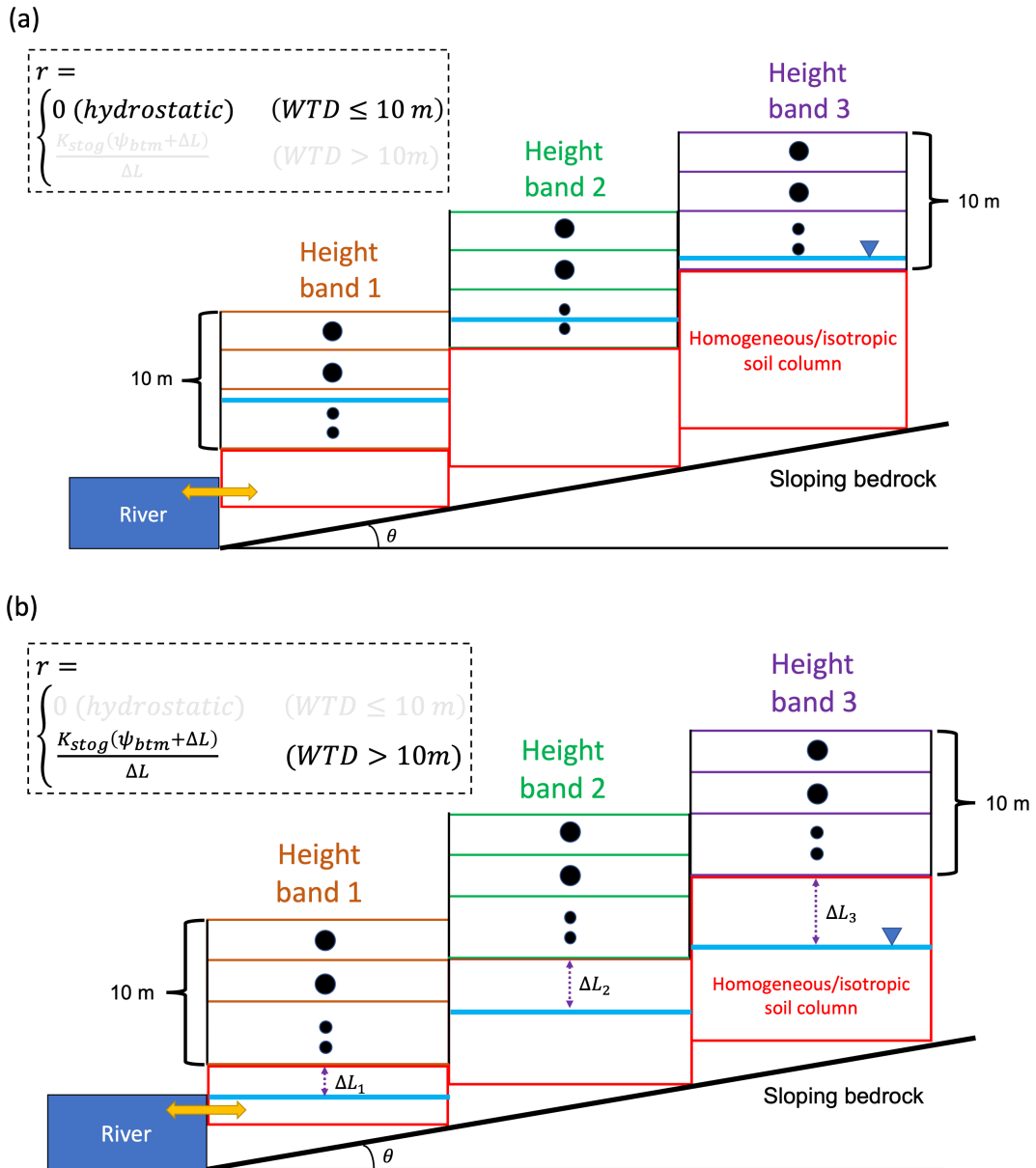
209   the vertical liquid flux (r in Equation 3) defines the soil columns' bottom boundary condition if the water table is

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

211   ground, the flow rate at the soil base is considered zero since the equipotential line is assumed to be vertical in the

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

213   allows the soil bottom drainage (SBD) at 10 m depth. This enables the consideration of the effects of groundwater

214   on the unsaturated soil processes, depending on groundwater configuration, such as groundwater properties and the

215   water table depth.

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

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

218   where r_i^j is the vertical water flux (mm s^{-1}) and the $K_{stog,i}^j$ is the hydraulic conductivity between the unsaturated soil

219   bottommost layer and the water table at i th height band at j th time step (mm s^{-1}), which is calculated by the harmonic

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

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

222   (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 hillslope length

223   from the reach of i th height band characterized by the HMC approach (m). ΔL_i^j is the distance between the soil

224   bottommost node and the water table at i th height band at j th time step (m).

225

(a)

$$r = \begin{cases} 0 \ (hydrostatic) & (WTD \leq 10 \ m) \\ \frac{K_{stog}(\psi_{btm}+\Delta L)}{\Delta L} & (WTD > 10m) \end{cases}$$

(b)

$$r = \begin{cases} 0 \ (hydrostatic) & (WTD \leq 10 \ m) \\ \frac{K_{stog}(\psi_{btm}+\Delta L)}{\Delta L} & (WTD > 10m) \end{cases}$$

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 was changed from zero-flux to variable-flux according to the hydraulic gradient between the bottommost soil layer and water table. WTD denotes water table depth.

8

234   **2.2.2. Two-way energy transport and conservation in LM4-SHARC**

235   LM4-SHARC accounts for the phase change of water in the groundwater domain by calculating the ice

236   content according to the groundwater temperature (Equation 9).

237
$$ws_{gw,i}^{j+1} = \begin{cases} ws_{gw,i}^{j} - \min\left(ws_{gw,i}^{j}\left(T_{gw,i}^{j} - T_{freeze}\right)\dfrac{hc_{gw,i}^{j}}{hf}\right) & (ws_{gw,i}^{j} > 0\,,\, T_{gw,i}^{j} > T_{freeze}) \\[2ex] ws_{gw,i}^{j} + \min\left(wl_{gw,i}^{j},\left(T_{freeze} - T_{gw,i}^{j}\right)\dfrac{hc_{gw,i}^{j}}{hf}\right) & (wl_{gw,i}^{j} > 0,\ T_{gw,i}^{j} < T_{freeze}) \end{cases} \tag{9}$$

238   where wl_{gw}, ws_{gw} are the liquid and ice content in the groundwater, respectively (-). T_{gw} is the groundwater

239   temperature (K), and T_{freeze} the freezing point of 273.15 K. hf is the latent heat of fusion, a constant of $3.3358 \cdot 10^5$

240   J kg^{-1}. hc_{gw} is the groundwater heat capacity, calculated by Equation 10.

241    $$hc_{gw,i}^{j} = hc_{btm,i}^{j} \times \left(e_{btm,i} - hl_i \tan\theta\right) + clw \times wl_{gw,i}^{j} + csw \times ws_{gw,i}^{j} \tag{10}$$

242   where $hc_{btm,i}^{j}$ the dry soil heat capacity of the bottommost soil layer at i th height band at j th time step (J K^{-1} m^{-2}).

243   clw is the specific heat of liquid (4218.0 J kg^{-1} K^{-1}), and csw is the specific heat of ice (2106.0 J kg^{-1} K^{-1}). We also

244   note that each HB-scale groundwater domain's thermal properties were assumed to be identical to those of the

245   corresponding soil column's bottommost soil layer.

246   The groundwater temperature in each HB is dynamically updated, taking into account time-dependent heat

247   capacity and heat fluxes conducted and advected from/to adjacent flow domains. Similar to the water table-

248   dependent boundary condition at the soil base, the heat flux boundary condition at the soil base is also affected by

249   the groundwater condition since the water advection is zero if the water table depth is less than 10 m (Equation 11,

250   12, and 13).

251    $$\delta_{v_i}^{j} = \begin{cases} \left(\delta_{v_{adv,i}}^{j} + \delta_{v_{cnd,i}}^{j}\right) & (water\ table\ depth > 10\ m) \\[1ex] \delta_{v_{cnd,i}}^{j} & (water\ table\ depth \leq 10\ m) \end{cases} \tag{11}$$

252    $$\delta_{v_{cnd,i}}^{j} = \lambda_{stog,i}^{j}\left(T_{gw,i}^{j} - T_{btm,i}^{j}\right) \tag{12}$$

253    $$\delta_{v_{adv,i}}^{j} = \begin{cases} \dfrac{r_i^{j}}{\rho}\left(T_{btm,i}^{j} - T_{freeze}\right)\dfrac{hc_{gw,i}^{j}}{\Delta t \Delta L_i^{j}} & (r > 0) \\[2ex] \dfrac{r_i^{j}}{\rho}\left(T_{gw,i}^{j} - T_{freeze}\right)\dfrac{hc_{gw,i}^{j}}{\Delta t \Delta L_i^{j}} & (r \leq 0) \end{cases} \tag{13}$$

254   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

255   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

256   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

257   layer and groundwater. The direction for $\delta_{v_{adv,i}}^j$ is determined by the water flux (i.e., downward recharge or upward

258   capillary) according to the hydraulic gradient. Thus, the soil columns and groundwater temperatures are simulated

259   with the modified heat flux boundary condition (at the soil column base) from constant thermal to variable-thermal

260   fluxes in the LM4-SHARC. The lateral heat transport in the groundwater domain ($\delta_{l_i}^j$) is another component in

261   determining the groundwater temperature (Equations 14, 15, and 16).

262    $$\delta_{l_i}^j = -q_{l,i}^j clw\left(T_{gw,i}^j - T_{freeze}\right) \qquad\qquad (i = n_{HB}) \qquad\qquad (14)$$

263    $$\delta_{l_i}^j = clw\left[q_{l,i+1}^j\left(T_{gw,i+1}^j - T_{freeze}\right) - q_{l,i}^j\left(T_{gw,i}^j - T_{freeze}\right)\right] \qquad (i = 2,..,n_{HB}-1) \qquad (15)$$

264    $$\delta_{l_1}^j = \begin{cases} clw\left[q_{l,2}^j\left(T_{gw,2}^j - T_{freeze}\right) - q_{l,1}^j\left(T_{gw,1}^j - T_{freeze}\right)\right] & (q_l \geq 0) \\ clw\left[q_{l,2}^j\left(T_{gw,2}^j - T_{freeze}\right) - q_{l,1}^j\left(T_{stream}^j - T_{freeze}\right)\right] & (q_l < 0) \end{cases} \qquad (16)$$

265   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

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

267   height band at j th time step (Equation 17).

268    $$T_{gw,i}^{j+1} = T_{gw,i}^j - \frac{\left[\left(\delta_{v_{cnd,i}}^j - \delta_{v_{adv,i}}^j\right) + \delta_{l_i}^j\right]\Delta t}{hc_{gw,i}^j} \qquad\qquad (17)$$

269   For the lateral heat exchange fluxes between the stream and HB$_1$ ($\delta_{l_1}^j$ in Equation 16), the stream

270   temperature is considered if the channel loses water to the riparian zone (i.e., $q_l < 0$). The stream temperature is

271   estimated considering how much heat flows into the stream from the hillslopes and out of it through the catchment

272   outlet (Equation 18).

273    $$T_{stream}^{j+1} = T_{stream}^j + \frac{\left[\left(\delta_{l_1}^j + \delta_{unsat\,i}^j\right) - clw\left(T_{stream}^j - T_{freeze}\right)\frac{\rho Q^j}{A^j}\right]\Delta t}{hc_{stream}^j} \qquad\qquad (18)$$

274   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

275   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

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

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

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

279

280   **2.3. The streamflow recession analysis for the groundwater parameterization**

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

282   as a function of streamflow Q (Equation 19). Since the analytical solutions to the Boussinesq equation can be recast

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

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

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

286   and Mohanty, 2023b).

287     $$-\frac{dQ}{dt} = aQ^b \hspace{4cm} (19)$$

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

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

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

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

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

293   effective groundwater properties.

294

295   **2.3.1. Selecting analytical models**

296   Theoretical catchment outflow from the Boussinesq groundwater yields two hydraulic regimes: the early

297   (i.e., high flow) and late time (i.e., low flow) domains. Since the LM4-SHARC considers sloping groundwater, we

298   only consider the analytical solution of the Boussinesq equation for a sloping groundwater. We used the analytical

299   solutions obtained by Brutsaert (1994), considering their applicability for a broader range of bedrock slopes

300   (Pauritsch et al., 2015). Thus, the recession parameter a for the early time domain is expressed in Equation 20 with

301   b_{early} being set to 3.0, and the parameter a for the late time domain is defined in Equation 21 with b_{late} being set to

302   1.0.

303     $$a_{early} = \frac{1.33}{K_s f T_{ini}^3 L^2 cos\theta}, \qquad b_{early} = 3.0 \hspace{3cm} (20)$$

304     $$a_{late} = \frac{\pi^2 p K_s T_{ini} L^2}{f A^2} cos\theta \left[1 + \left(\frac{\frac{B}{T_{ini}} tan\theta}{\pi p} \right)^2 \right], \qquad b_{late} = 1.0 \hspace{2cm} (21)$$

305   where A is the subsurface drainage area (m^2) that effectively contributes to the recession slope characteristics, and p

306   is a constant set to 1/3. B is the contributing groundwater's characteristic length (m) under the geometric similarity

307   assumption, calculated by $B=A/(2L)$.

308

309   **2.3.2. Event-scale recession analysis**

310   This study recognized that the recession parameters from the point cloud data (i.e., collective recession

311   data) could be artifacts of the variability in individual recession events (REs) (Jachens et al., 2020; Tashie et al.,

312   2020; Karlsen et al., 2019). To fill this gap, we performed an event-scale recession analysis to account for the

313   variability of recession slope characteristics among individual REs, which results in different estimates of

314   groundwater properties. The onset of the REs was starting 5 days after the peak to exclude the influence of overland

315   flow (i.e., runoff) on streamflow hydrograph. We examined the consecutive decline of daily discharge observational

316   data to decide the duration of an RE, and the end date of an RE was determined when the daily stream discharge was

317   at its lowest.

318   For each RE, we performed the logarithmic regression of Q on $-\frac{dQ}{dt}$ in bi-logarithmic space. The time

319   derivative of Q ($\frac{dQ}{dt}$) was computed based on the daily streamflow. Each RE's transition point of the hydraulic regime

320   from the early to the late time domain was identified daily through the method suggested by Hong and Mohanty

321   (2023a). In this method, the transition point from the early time to the late time domain is determined based on an

322   abrupt and most noticeable change in R^2 values from linear regression with a fixed slope of 3.0 while incrementing

323   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

324   selectively used the daily streamflow time series of REs that lasted for more than 20 days to get enough data pairs of

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

326   regime transition in an event-scale recession analysis, see Hong and Mohanty (2023a).

327

328   **2.4. Characterization of the catchment-scale domain and its heterogeneity in LM4-SHARC**

329   A hierarchical multivariate clustering (HMC) approach (Chaney et al., 2018; Chaney et al., 2016) was used

330   to characterize the catchment-scale model domain and its spatial heterogeneity. The conceptual approach uses

331   available environmental datasets (such as soil properties, topography, meteorology, and land cover) to characterize

332   the sub-grid spatial heterogeneity within each grid cell of an Earth System Model (ESM) land domain (Chaney et

333   al., 2018). In fact, the HMC method has been previously used in constructing land domains for LM4-HydroBlocks

334   to account for the sub-grid spatial variability of land properties within a regular latitude-longitude climate grid cell

335   (e.g., at 0.5° or 1.0° resolution). The land fraction of each grid cell in the LM4-HydroBlocks domain is partitioned

336   into soil, glacier, and lake components. The soil component of a grid cell is composed of hillslopes clustered in a set

337   of k characteristic hillslopes (CH). Each CH has unique attributes, such as slope, aspect, and convexity, which are

338   local averages of these fields obtained by high-resolution datasets. Each CH is further partitioned into l units

339   denoted as "height bands" (HB), which are obtained by partitioning each CH based on elevation bins of size dh.

340   Finally, each HB can be further divided into p clusters (i.e., tiles), and then each tile is considered representative

341   element volume (REV) of soil with respect to covariates incorporated for clustering (Chaney et al., 2018). The

342   channels were delineated using an area threshold of 100,000 m^2. Three parameters k, dh, and p are required in the

343   land pre-processing code; k determines the number of CH in the land model domain, dh defines the height difference

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

345   HB (i.e., the intra-HB variability). For further details on the HMC algorithm, see Chaney et al. (2016); Chaney et al.

346   (2021).

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

348   time, extended the existing HMC framework to generate a land input dataset for a catchment domain with an

349   irregular boundary. This refinement has been designed to assess the GFDL land model's performance at the

350   catchment level while enhancing the interpretability of the land model's hydrologic outputs. The catchment's

351   boundary was determined by computing the geographic extent of the contributing area that drains to a specific point

352   (i.e., the catchment outlet). Since the catchment areas are significantly variable (roughly ranging from sub-km^2 to

353   hundreds of km^2), the resolution of a digital elevation model (DEM) used for clustering the terrain characteristics

354   should be fine enough to capture the intra-catchment heterogeneity of terrain characteristics adequately. To this end,

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

356

2.5. Experimental design for model comparison

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

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

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

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

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

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

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

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

366   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

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

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

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

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

371   precipitation and air temperature were assimilated into the GSWP3 forcing data by being directly inserted (i.e.,

372   direct insertion data assimilation). This was to improve the accuracy of the model outputs, given the inconsistency

373   between the GSWP3 forcing and the local meteorological conditions primarily due to scale difference (i.e.,

374   0.5°×0.5° vs. 1 km^2). We considered precipitation and air temperature the most significant atmospheric variables

375   determining the catchment's water and energy budgets, so that the in-situ precipitation and air temperature data

376   obtained during the evaluation period replaced the corresponding variables in the GSWP3 forcing. LM4-

377   HydroBlocks and the LM4-SHARC were then operated using the identical forcing input. Two statistical metrics of

378   Pearson's correlation coefficient R^2 and $RMSE$ (Root Mean Square Error) were used to evaluate the temporal

379   agreement of the modeled hydrologic outputs against corresponding observations and the errors, respectively.

380

3. Study area and observational data

3.1. Study area and the HMC parameters

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

384 Nevada (Figure 2 a). The P301 headwater catchment is one of eight primary headwater catchments of the Kings

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

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

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

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

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

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

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

392 developed LM4-SHARC model.

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

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

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

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

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

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

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

400 of rain and snow, and precipitation transitions from majority rainfall to majority snow approximately at 2000 m in

401 elevation (Bales et al., 2018; Hunsaker et al., 2012).

402   In compiling the land input dataset for the given catchment, the *k* was set to 1 as the study area is a single

403 headwater catchment. The surface elevation data from 3DEP DEM was used as the sole variable to account for the

404 intra-catchment variability of terrain properties, and each HB's extent was determined using a *dh* value of 20 m. *p*

405 was set to 1; however, we note that the number of tiles in each HB can increase (or decrease) depending on factors

406 such as natural mortality, land use, and fire events applied to each tile. The stream flowline was also delineated

407 according to the 3DEP DEM using the area threshold of 100,000 m^2. Consequently, the P301 catchment was

408 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

409 field measurements (Figures 2 b, c, and d).

410
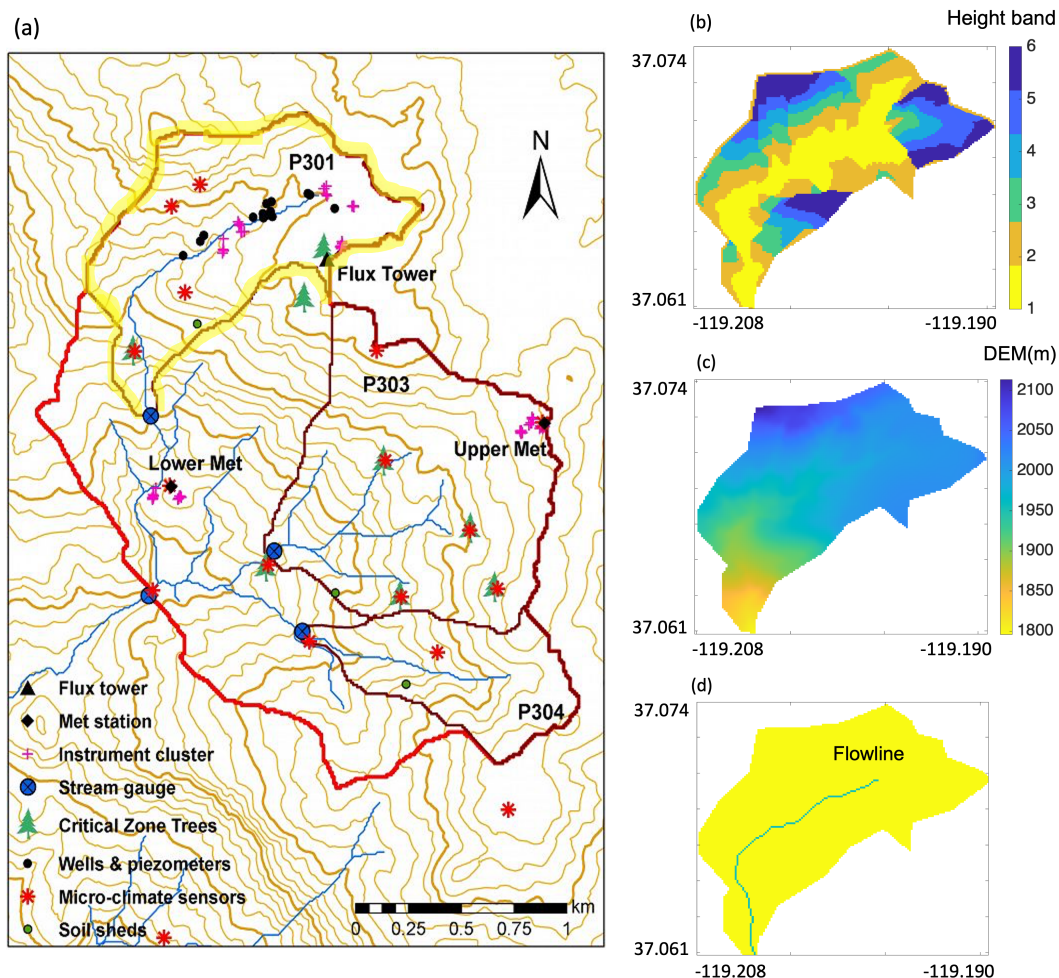
Figure 2. (a) The yellow highlighted area indicate the spatial extent of the headwater catchment P301. While the soil moisture and temperature data measured at the Lower Met station is not within the P301 catchment, the measurements from the Lower Met station was used due to its proximity to the catchment 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.

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

426   discharge rate using the standard rating curve supplied by the flume and weir manufacturers (Bales et al., 2018). The

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

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

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

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

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

432   separation method ensures consistency between the applied analytical/numerical models and the observational

433   baseflow data.

434

435   **3.2.2. Soil moisture, soil temperature, and snow depth**

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

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

438   60, and 90 cm below the soil surface, and the measurements were used as representative values for each soil depth

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

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

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

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

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

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

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

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

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

448   2009 – WY 2012 (4 years). The SMC and ST observations measured at the depth nearest to the land surface (10 cm)

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

450

451   **3.2.3. Meteorological observations**

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

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

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

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

456

457   **4.   Results and discussion**

458   **4.1.   Suitability of the select analytical model**

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

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

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

462   (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

463   (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

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

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

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

467   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

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

469   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

470   simulations.

471   The temporal agreement and the total magnitude of groundwater divergence fluxes per unit width (i.e., $q_l N$)

472   were investigated during the recession duration of 15 days. As shown in Figure 3, $q_l N$ decreases as the hydraulic

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

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

475   (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

476   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

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

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

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

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

481   to tune the numerical Boussinesq groundwater with extremely high D and steeper bedrock (Figure 3 c). Except for

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

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

(a)

(b)

(c)
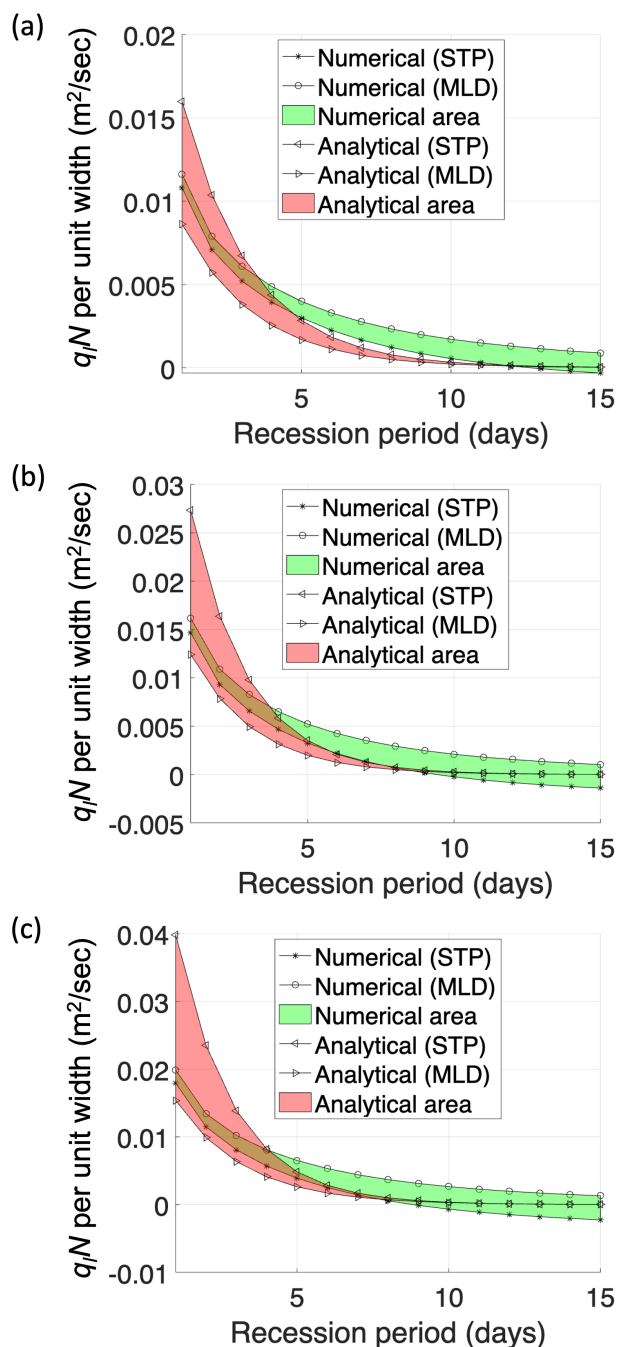
Figure 3. Comparison between the respective baseflow flux solutions derived from the selected analytical and numerical models.
(a) the SLW case with K_s of 2.0E-03 mm/sec and f of 0.20, (b) the NRM case with K_s of 9.0E-03 mm/sec and f of 0.05, and (c)
the FST case with K_s of 2.0E-02 mm/sec and f of 0.02.

489   **4.2. Event-scale recession analysis and calibration**

490   **4.2.1. Quantifying the uncertainty of hydraulic diffusivity estimates**

491   We performed the event-scale recession analysis to estimate the effective groundwater properties of the

492   study catchment. Following the extraction criteria (section 3.3.2), 18 individual REs were extracted from the

493   streamflow observations of 13 years, and their recession parameters a and b were estimated for each hydraulic

494   regime (Table 1). Figure 4 (a) presents an example showing how an individual RE was analyzed with the selected

495   analytical models for its early-time and late-time domains using daily streamflow data from the P301 catchment.

496   Once the recession parameters a and b were estimated for each RE, the range of bedrock slope θ should be

497   adequately constrained in order to determine hydraulic diffusivity parameters K_s and f. The characteristic length B is

498   calculated by assuming the geometric similarity between both hillslope sides in the catchment; thus, $B=A/(2L)$. The

499   initial saturated groundwater thickness N_{ini} is considered constant across the individual REs, equivalent to the initial

500   value of groundwater thickness applied in the LM4-SHARC (here, 10 m). Also, the entire catchment (i.e., 1 km^2)

501   was assumed to contribute to the hydrograph recession characteristics at the catchment outlet. The following criteria

502   were applied to define the upper and lower bound of θ: (1) the effective porosity f should range from 0.1% to 20.0%

503   (Brutsaert and Lopez, 1998; Tashie et al., 2021; Troch et al., 1993; Hong and Mohanty, 2023b; Hong and Mohanty,

504   2023a; Heath, 2004), and (2) the catchment-scale effective groundwater lateral hydraulic conductivity K_s cannot

505   exceed 1.0 mm s^{-1} (Fan and Miguez-Macho, 2011; Gómez-Hernández and Gorelick, 1989; Tashie et al., 2021).

506   The variability of recession characteristics was quantified by the recession parameter a (since b is fixed) in

507   the early and late time domains (i.e., a_{early} and a_{late}) (Figure 4 b). Essentially, this parameter provides insights into

508   the variability of the groundwater's effective properties dependent on the memory effects of the catchment (e.g.,

509   groundwater storage). The 98% confidence intervals for a_{early} and a_{late} were estimated to be [-3.10, -2.43] and [-6.81,

510   -6.20] respectively from the analysis of 18 REs. Following the above criteria, the value distributions of K_s, f, and θ

511   corresponding to the ranges of a_{early} and a_{late} were estimated. The uncertainty of K_s, f, and θ was then quantified by

512   determining the intersection range between K_s, f, and θ values derived from the lowest a (i.e., $\log\left(-\frac{dQ}{dt}\right) =$

513    $-3.10 \log(dQ) - 6.81$) and the highest a (i.e., $\log\left(-\frac{dQ}{dt}\right) = -2.43 \log(dQ) - 6.20$). The upper and lower bounds

514   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

515   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,

516   the relationship between hydraulic diffusivity (per unit wetted perimeter) D and the bedrock slope θ follows a power

517   function. Consequently, we considered that the (θ, D) space in which solutions of D and θ (potentially representing

518   the catchment average behavior) could exist was further constrained by the specified range of parameters K_s, f, and θ

519   between the power functions (Figure 4 b).

520

521

522

523

524

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	3.0	-6.76	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	

Table 1. Recession period, recession characteristics, and parameters *a* and b for each recession event under different diffusivity conditions. The variability of the parameter a indicates the variability of distinct diffusivity of groundwater across individual recession events.
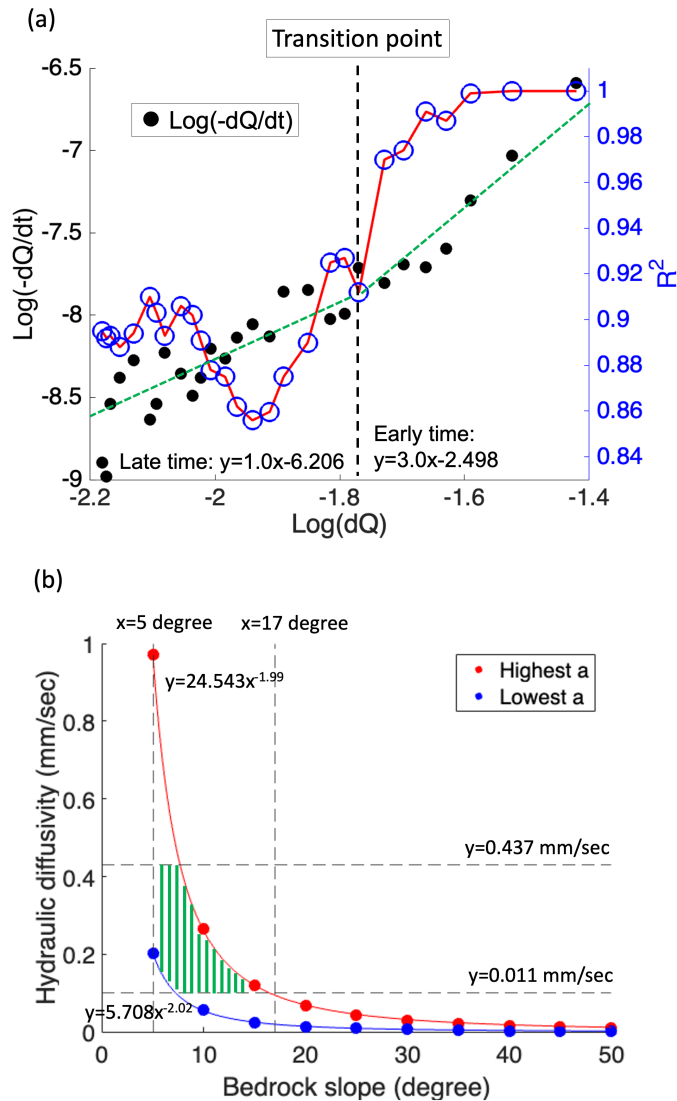
Figure 4. (a) An example showing how the transition point of hydraulic regime (from early-time to late-time domain) is determined from an individual recession event. The combined understanding of the recession parameter a_{early}, and a_{late} enables inferring the groundwater properties such as hydraulic 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

21

539　LM4-SHARC's groundwater domain. Based on the identified value range of θ and D, we tried to determine a single

540　pair of (θ, D) that shows the optimal accuracy by comparing the modeled and observed baseflow data. The modeled

541　baseflow fluxes were estimated by summing the liquid fluxes from saturated soil to the stream, and baseflow

542　observations during this study's evaluation period (from 10/1/2003 – 9/30/2014, WY2004 - 2014) were used for

543　calibration. We considered a (θ, D) that best represents the temporal dynamics and magnitudes of baseflow

544　observations as the calibrated (θ, D) for the study catchment.

545　　　We identified that a specific parameter pair (θ, D) can be specified in the uncertainty space. Figure 5 shows

546　that (1) the temporal agreements between the modeled and observed baseflow were generally related to the

547　magnitude predictions, (2) the most significant improvements of R^2 and $RMSE$ (from the LM4-HydroBlocks to

548　LM4-SHARC) were identified coincidentally at a specific pair of (θ, D). We found the most pronounced

549　improvements of baseflow predictions in the LM4-SHARC compared to the LM4-HydroBlocks, with the R^2

550　improvement of 0.155 and $RMSE$ reduction of 0.220 mm d^{-1} at θ=5.0 degree, D=4.6$\cdot$10^{-2} mm s^{-1}. The R^2 and

551　$RMSE$ of the LM4-SHARC-derived baseflow estimates against the observations (over the 11 years) were estimated

552　at 0.402 and 0.556 mm d^{-1}, respectively, while those of LM4-derived baseflow estimates were estimated at 0.247

553　and 0.776 mm d^{-1}, respectively. With the calibrated groundwater properties, we confirmed the recession behaviors of

554　the P301's streamflow hydrograph over the 11-year evaluation period were generally better captured in the LM4-

555　SHARC's baseflow estimates compared to those of the LM4-HydroBlocks (Figure 6).

(a)

(b)
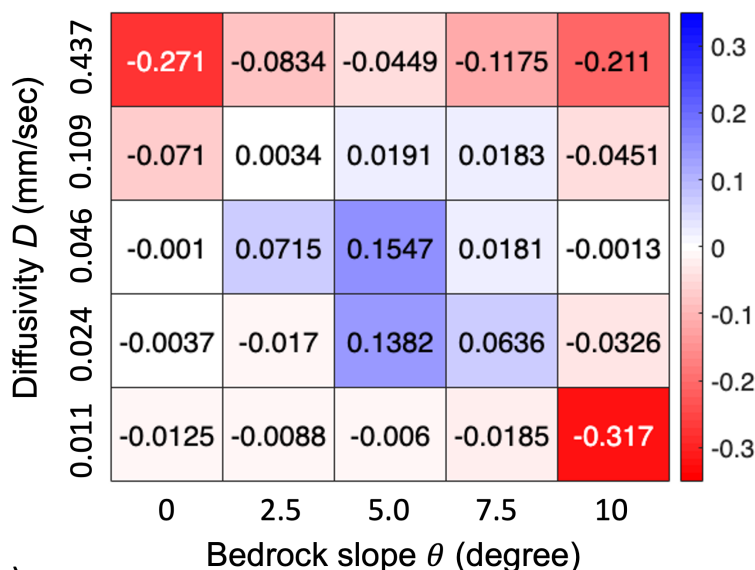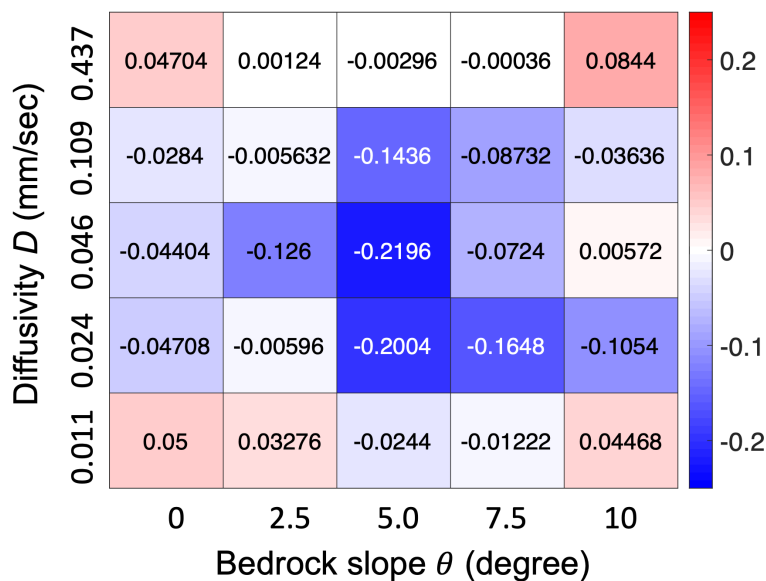
556
557 **Figure 5**. The pair of groundwater parameters (θ, D) was specified to calibrate the catchment groundwater. (a) R^2 difference
558 (improvement) between the LM4-SHARC and LM4-HydroBlocks ($R^2_{LM4\text{-}SHARC}$ - $R^2_{LM4\text{-}HydroBlocks}$). Here, R^2 denotes the coefficient
559 of determination between the modeled baseflow and observations for the 11 years evaluation period, (b) for the same period,
560 $RMSE$ difference (reduction) between the LM4-SHARC and LM4- HydroBlocks ($RMSE_{LM4\text{-}SHARC}$ - $RMSE_{LM4\text{-}HydroBlocks}$) was
561 evaluated.
562
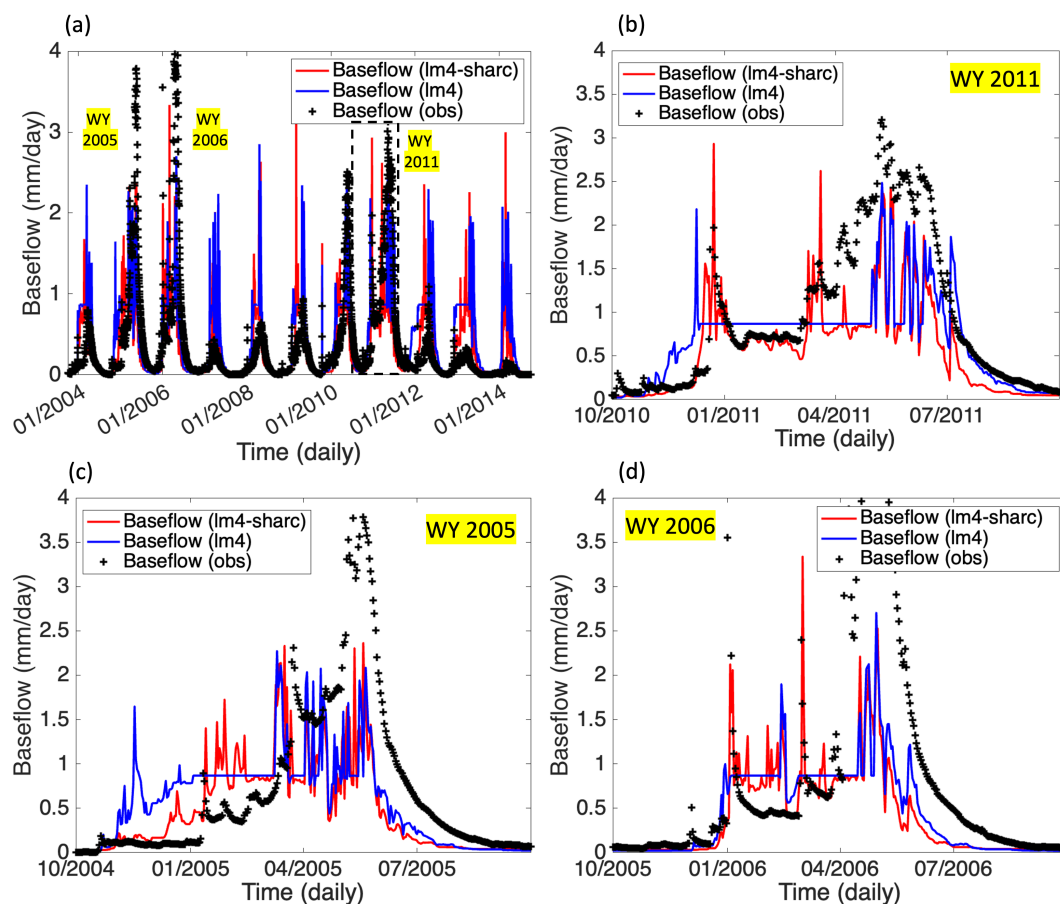
Figure 6. Comparative time-series of baseflow estimates from the LM4-SHARC (red), LM4-HydroBlocks (indicated LM4 in the legend, blue), and the corresponding observations. (a) Daily time-series of the baseflow data over the evaluation period of 11 years, (b) time-series in Water Year (WY) 2011, (c) time-series 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

Based on the improved baseflow estimates, we assessed the effects of groundwater-induced soil processes on the near-surface water and energy budgets. We applied $\theta=5.0$ degree, $D=4.6\cdot 10^{-2}$ 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 was performed for four years, from WY 2009-2012. Figures 7 (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

24

579   period, we identified that soil bottom drainage (SBD) facilitation due to the modified boundary condition from zero-

580   flux to variable-flux BC at soil columns' base (Figure 1 a) significantly affected the soil moisture content (SMC) at

581   10 cm depth. The time series of 10 cm SMC from the LM4-SHARC showed significantly reduced SMC compared

582   to LM4-HydroBlocks. The entire soil columns reached full saturation too quickly in the LM4-HydroBlocks, and the

583   facilitated downward liquid transport due to drier soil lower layers in the LM4-SHARC was found to be effective in

584   correcting the soil's wet biases. As a result, the temporal variability and total storage of the observed SMC was

585   better captured in the LM4-SHARC. The average of the four R^2 from the yearly time series comparison (i.e., 2009-

586   12) was improved from 0.831 (LM4-HydroBlocks) to 0.849 (LM4-SHARC) while noticeably reducing the average

587    $RMSE$ from 0.0722 m^3 m^{-3} (LM4-HydroBlocks) to 0.0425 m^3 m^{-3} (Figure 8 a and b).

588   We also found that the enhanced representation of SMC resulted in better capturing of near-surface soil

589   temperature dynamics. The decreased SMC reduces the evapotranspiration rate, especially during daytime, leading

590   to increased sensible heat in the soil's energy balance. Also, the reduced soil heat capacity due to the decreased

591   SMC (i.e., liquid water) could increase the soil temperature under the given net radiation. Consequently, we

592   identified that soil temperature predictions at the 10 cm depth showed significant improvements in the LM4-

593   SHARC, primarily during warmer seasons, and showed better agreement with the observations. The soil temperature

594   estimates from both model configurations showed similar values when the surface was covered by snow (i.e., snow

595   depth > 0). In this case the soil is insulated by snow and thus variations in soil water predicted by the two model

596   configurations do not lead to appreciable differences in soil temperature. From the improved skill in predicting near-

597   surface soil temperature in LM4-SHARC, we conclude that the soil columns with applied zero-flux BC at 10 m

598   depth could hold too much soil water due to the imposed shallow water table. This overestimation in soil water

599   content could lead to an inaccurate description of land energy balance (e.g., overestimated ET/less sensible heat),

600   and thus to biases in soil temperature. We note that the average of the four R^2 from the yearly time series

601   comparison between the modeled and observed soil temperature increased from 0.952 to 0.957, and the $RMSE$

602   significantly reduced from 2.77 K to 1.67 K (Figure 8 c and d).

603   The simulated snow depth from both LM4-HydroBlocks and LM4-SHARC generally showed reasonable

604   agreement with the measured snow depth in the catchment. We note that the melt-out timing of snow involves a

605   mutual influence on soil temperature. Specifically, we noticed that the timing of snow melt-out is affected by soil

606   temperature, as the warmer ground expedites the melting. The melt-out timings of snow in the four evaluation years

607   were represented sooner in the LM4-SHARC than the LM4-HydroBlocks due to increased soil temperature (Figure

608   7). Also, as the snow melted out, the soil temperature quickly increased as the soil was no longer insulated by snow,

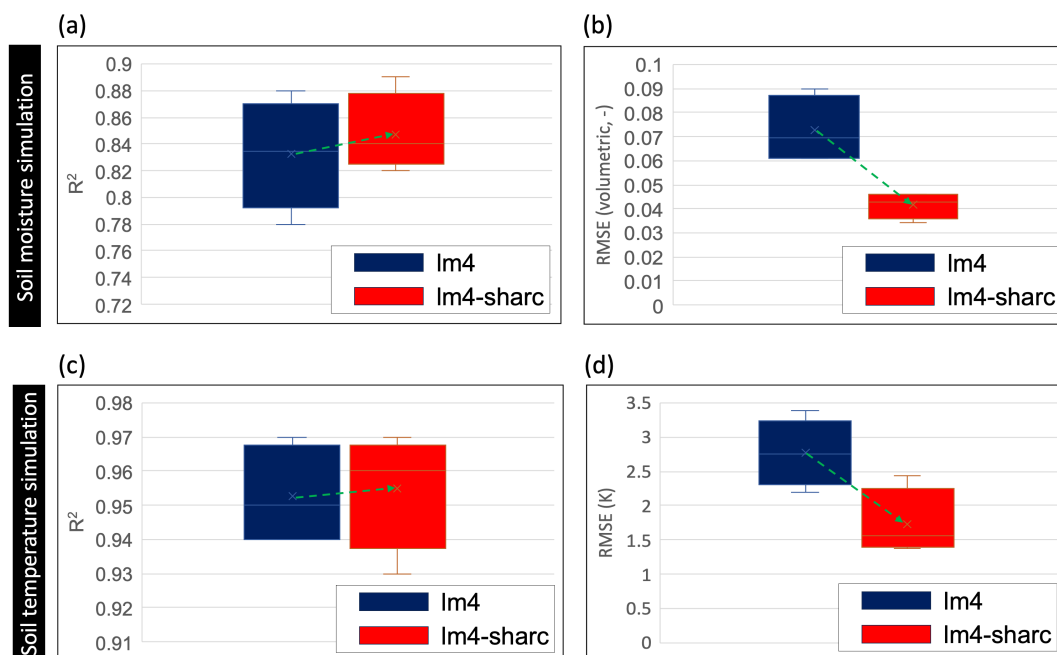609   leading to a higher correlation with surface air temperature.

610
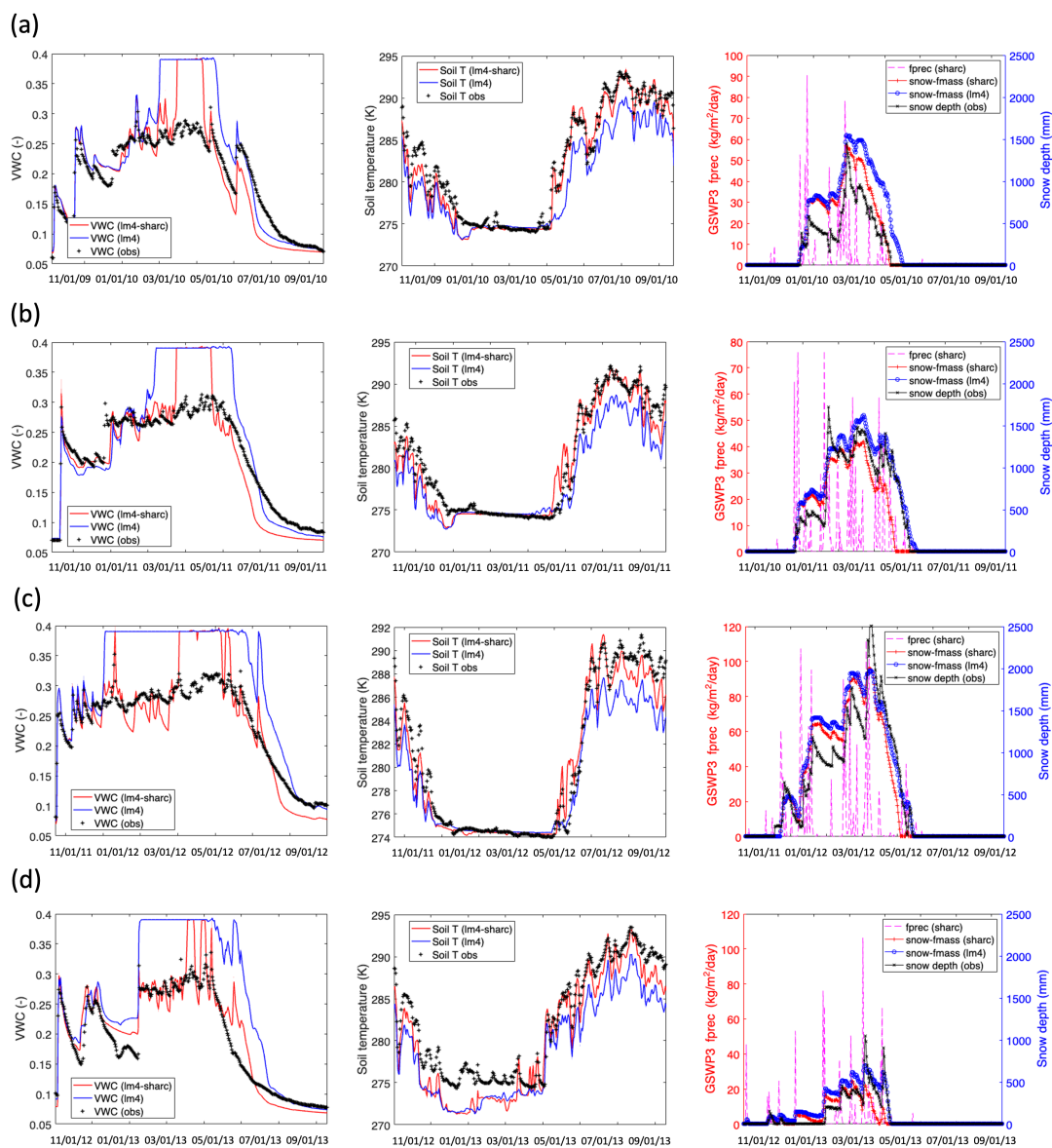
Figure 7. Comparative yearly time-series of soil moisture, temperature, and snow mass for the evaluation period from (a) WY2009, (b) WY2010, (c) WY2011, and (d) WY2012. For all the years, higher 10 cm depth soil temperature values due to drier soil were generally identified.

Figure 8. (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 LM-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 (section 3.4). Figure 9 (b), (c), (d), (e) 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). In addition to the three D conditions considered in Section 5.1, an additional consideration of a very slow flow D case (i.e., VSW, $K_s=1.0 \cdot 10^{-3}$ mm s^{-1}, f=0.2) was made to investigate the expected spin-up time when the lateral groundwater discharge flux is very slow. We found that the spin-up took less time (i.e., a smaller number of

27

640   10-year repeated cycles) as D increased, meaning that the LM4-SHARC reached a steady state more quickly with

641   the faster groundwater discharge. It turned out that the faster flux estimates could contribute more efficiently to the

642   convergence of other groundwater discharge-related variables (in the adjacent model domains), leading to a less

643   spin-up time. Likewise, from the comparison at the identical D condition, the slope θ can change the spin-up time

644   by affecting the flux velocity. This is because the q_l increases with increasing bedrock slope due to higher

645   quantitative growth rate of gravitationally-driven q_l component compared to its diffusion-driven component. As a

646   result, in our estimates the spin-up times at the VSW condition range from $60 - 120$ years, at the SLW condition 50

647    $- 110$ years, and at the NRM condition $40 - 80$ years. We found the spin-up time in FST conditions to be 30 years,

648   and to be independent of θ (Figure 9 a).

649   ΔSWS due to groundwater discharge was investigated based on the steady state of the model. The working

650   hypothesis here is that the lateral groundwater discharge could facilitate the SBD by inducing the downward

651   hydraulic gradients from the subsurface soil and water table. Figure 10 (a) shows the variability and magnitude of

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

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

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

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

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

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

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

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

660   greater than 0.1 mm s^{-1}, the lack of groundwater storage occurred irrespective of the bedrock slope.
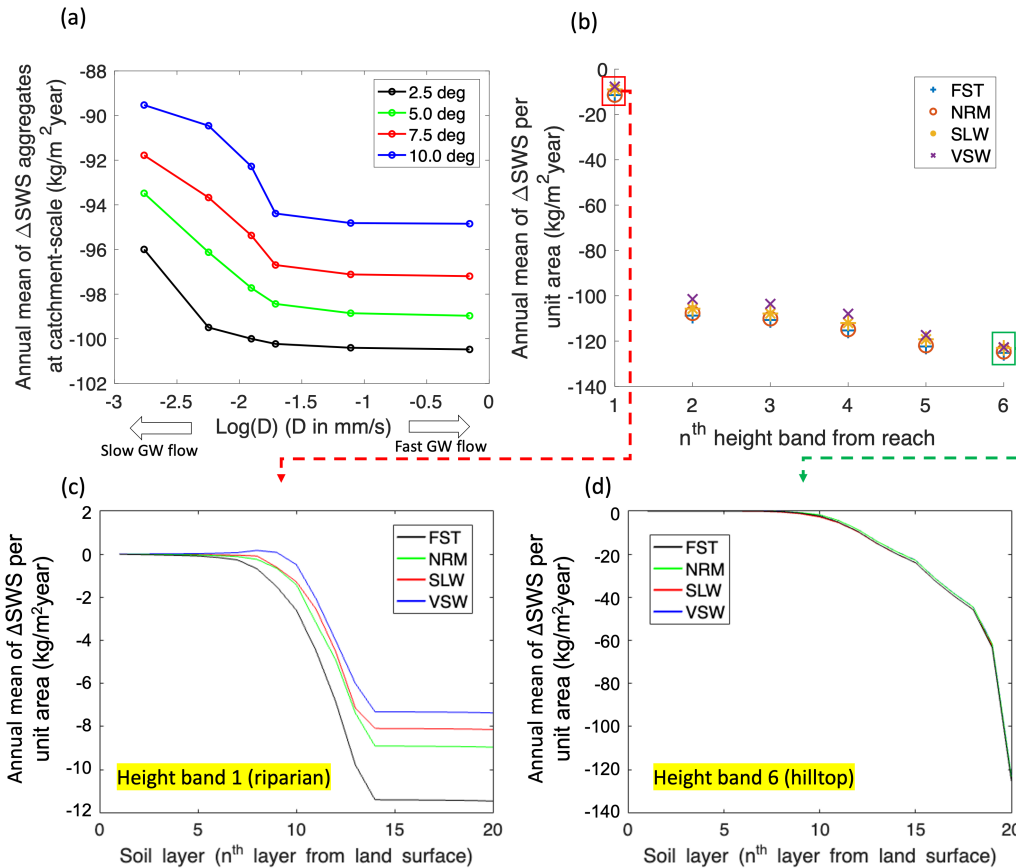
661   Moreover, the SBD facilitation due to the groundwater lateral discharge was more noticeable as the height

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

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

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

665   10 (c) and (d) illustrate that the effects of groundwater flow conditions on the SWS variability could be more

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

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

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

669   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

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

671   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),

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

673
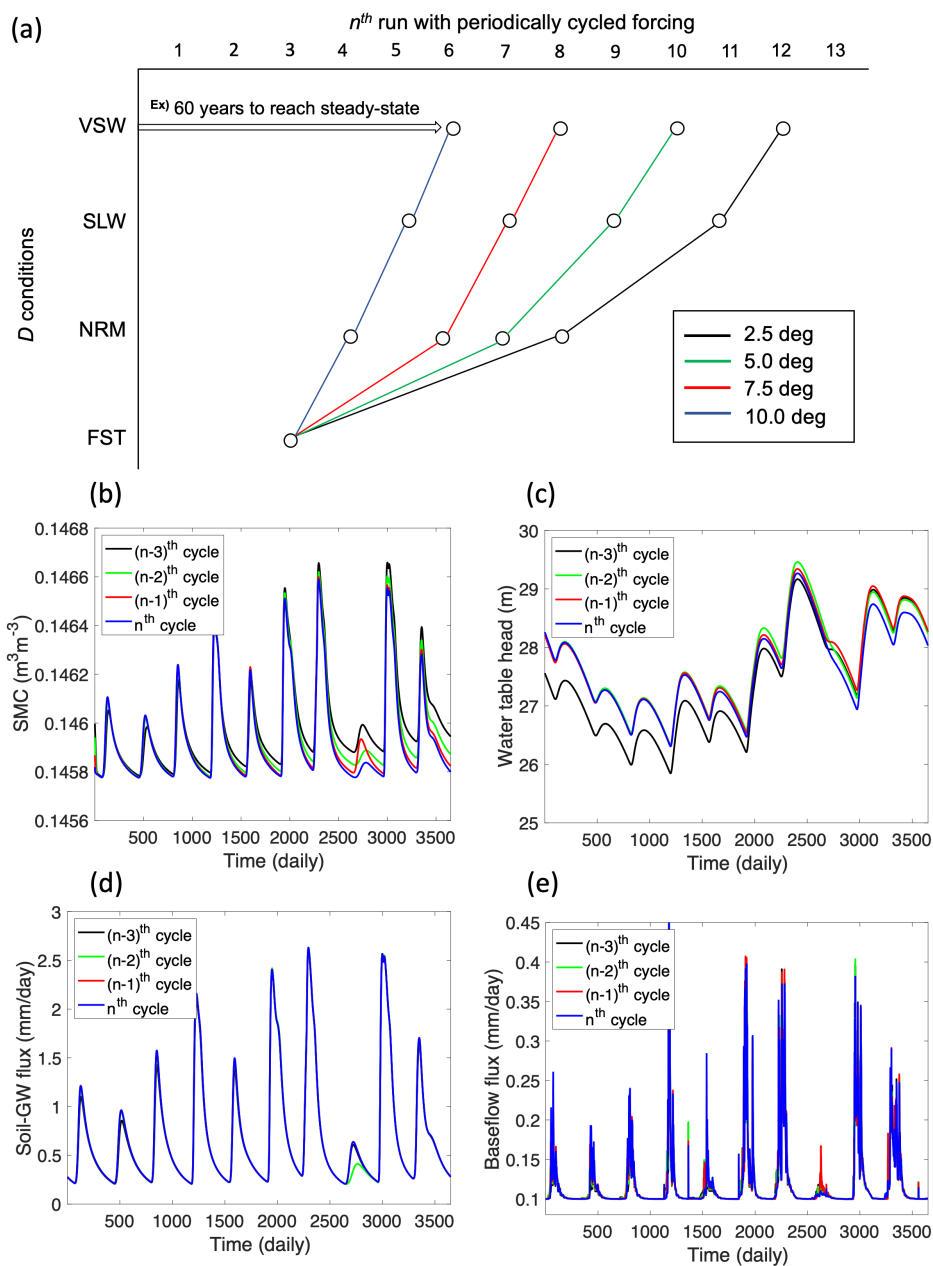
Figure 9. (a) varied spin-up time due to distinct groundwater properties θ, and D. The simulations of (b) the volumetric water content at the soil bottommost layer, (c) the groundwater table, (d) groundwater recharge fluxes (since positive), (e) baseflow flux reduce their deviations (from the previous cycle) as the cycles progress to be considered steady-state based on the differences from the nth cycle and $(n-1)^{th}$.

Figure 10. (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., HB6) were more dramatic than those at the riparian zone (i.e., HB1) as the vegetation evolved (Figure 11 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,

698   we found that the LAI contrast was more significant during the warmer season at the hilltop area, while the overall
699   trend of the LM4-SHARC-derived LAI evolved comparably to that of the LM4-derived LAI at the riparian zone
700   (Figure 11 c, and d). The hydrologic convergence at the riparian zone explains the comparable LAI dynamics at the
701   riparian zone as the subsoil domain readily saturated by the converging water impedes SBD, leading to higher water
702   retention in the partially saturated soil. Different soil moisture availability (between HB_1 and HB_6) resulting from
703   the contrasting SBD dynamics is thus emphasized by the density of plants, especially during the warm(er) season
704   when the plants yield higher transpiration rates. The varied transpiration rates in the LM4-SHARC also consistently
705   explain the LAI contrast due to different soil moisture conditions. We found that the transpiration rate at the hilltop
706   was reduced by 29.5% in the LM4-SHARC, while the rate was reduced by 10.3% in the riparian zone in the LM4-
707   SHARC (Figure 11 e and f). Overall, we found that the variations in SWS, transpiration, and LAI simulations are
708   consistent in that the groundwater convergence to the river valley intensified the catchment's contrasting hydrologic
709   states.
710
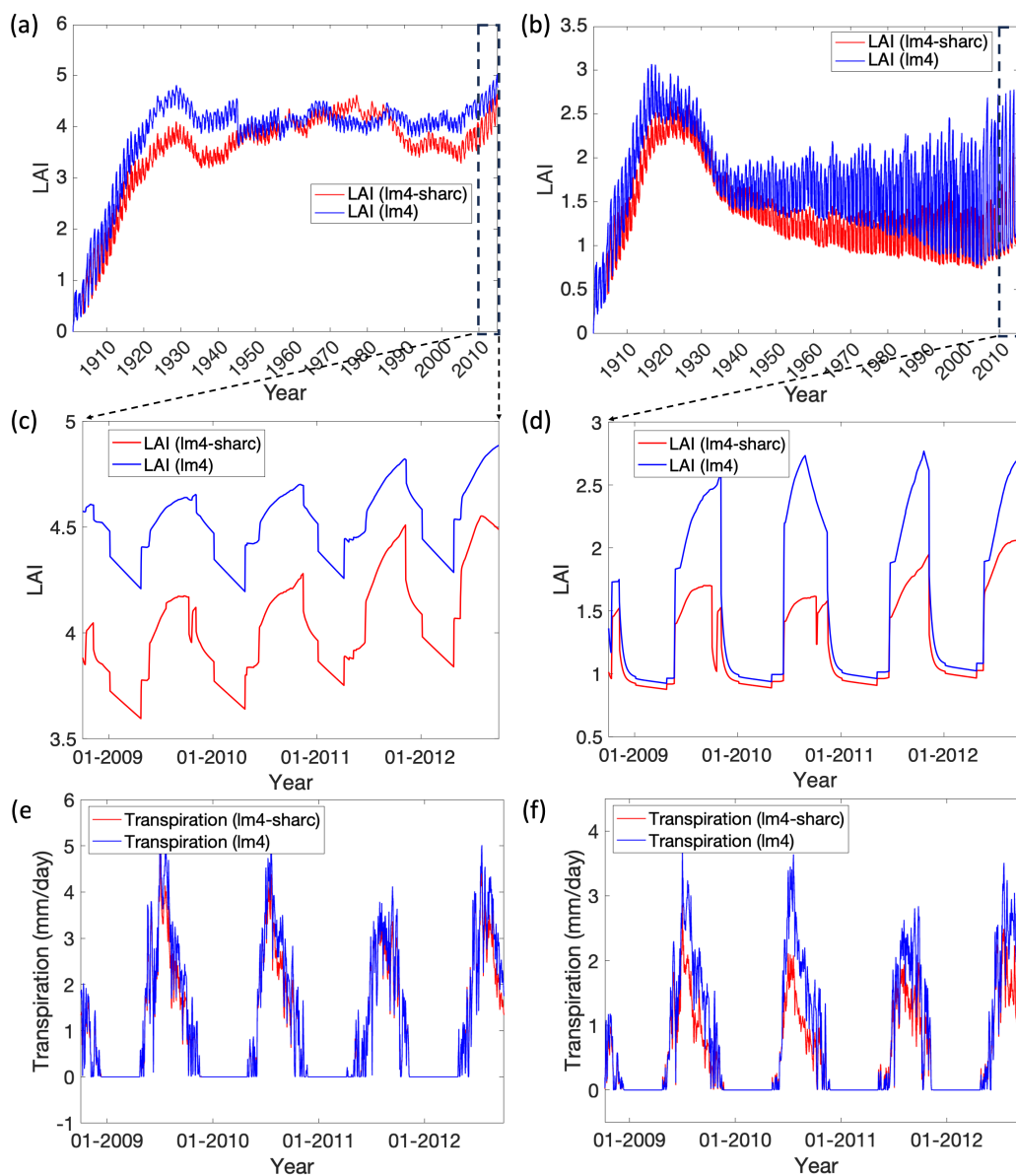
Figure 11. Temporal evolution of LAI from 1901 to 2014 (114 years) using GSWP3 forcing at (a) the riparian zone (HB1) and (b) the hilltop (HB6). The differences in the LAI (between the LM4-SHARC and LM4-HydroBlocks) at the hilltop area were more dramatic than those at the riparian zone. (c), (d) the 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.

720 **4.5. Applicability of the LM4-SHARC in an ESM**

721 **4.5.1. Testing LM4-SHARC in various climatic and orographic zones**

722   To support the implementation of LM4-SHARC in the GFDL ESM, we need to investigate the

723 performance of LM4-SHARC in various climatic and orographic zones. For this purpose, three additional headwater

724 catchments were selected based on their precipitation and topographic slope characteristics (Figure 12 a): the

725 Musselshell (MT), Maine (ME), and Clearwater (WA) headwater catchments. The precipitation and slope

726 characteristics of these sites vary from those of the Clearwater catchment (the wettest: 3,136 mm y^{-1}, and steepest:

727 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

728 this experiment, the groundwater properties in the additional catchments were assumed to be identical to those of the

729 P301 headwater catchment so that diffusivity D was set to 0.046 mm s^{-1} and the slope of groundwater bedrock θ was

730 assumed to be 5 degrees in all the three headwater catchments. The steady state of the groundwater and any adjacent

731 flow domains was also ensured, and the evaluation was performed using the 10-year model outputs.

732   We tried to identify if the model is robust under diverse conditions by examining the consistency between

733 hydrologic characteristics and indices. The hydrologic indices include: (1) runoff ratio (i.e., ratio of streamflow to

734 precipitation), (2) baseflow coefficient (i.e., ratio of baseflow to precipitation). The annual mean

735 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

736 281 mm y^{-1}/1,897.4 mm y^{-1} from the Musselshell, Maine, and Clearwater catchment, respectively. With the

737 respective annual mean precipitation of each catchment (i.e., Musselshell - 397.2 mm y^{-1}, Maine - 1223.1 mm y^{-1},

738 and Clearwater - 3136.6 mm y^{-1}), each catchment's baseflow coefficient/runoff ratio was estimated at 0.028/0.125,

739 0.086/0.298, and 0.230/0.604 the Musselshell, Maine, and Clearwater catchment, respectively. The established

740 positive correlation between the baseflow coefficient and runoff ratio is consistent with what is reported in existing

741 studies (Cheng et al., 2021; Ouyang et al., 2018) (Table 2). We also found that the value gradients of

742 baseflow/recharges estimates correspond to what is expected from the slope/precipitation gradients. For example,

743 the highest yield of baseflow from the Clearwater catchment (i.e., 0.77 mm d^{-1}) can be explained by its high

744 precipitation and steep slope, which contribute to higher drought flow (i.e., baseflow during dry seasons) and peaks.

745 Also, the partially saturated soil found in most parts of the Musselshell catchment explains the minimal baseflow

746 amounts due to the lack of groundwater storage from the limited groundwater recharge (Figure 12 b and c).

747

748

749

750

751

752

753

754

		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.
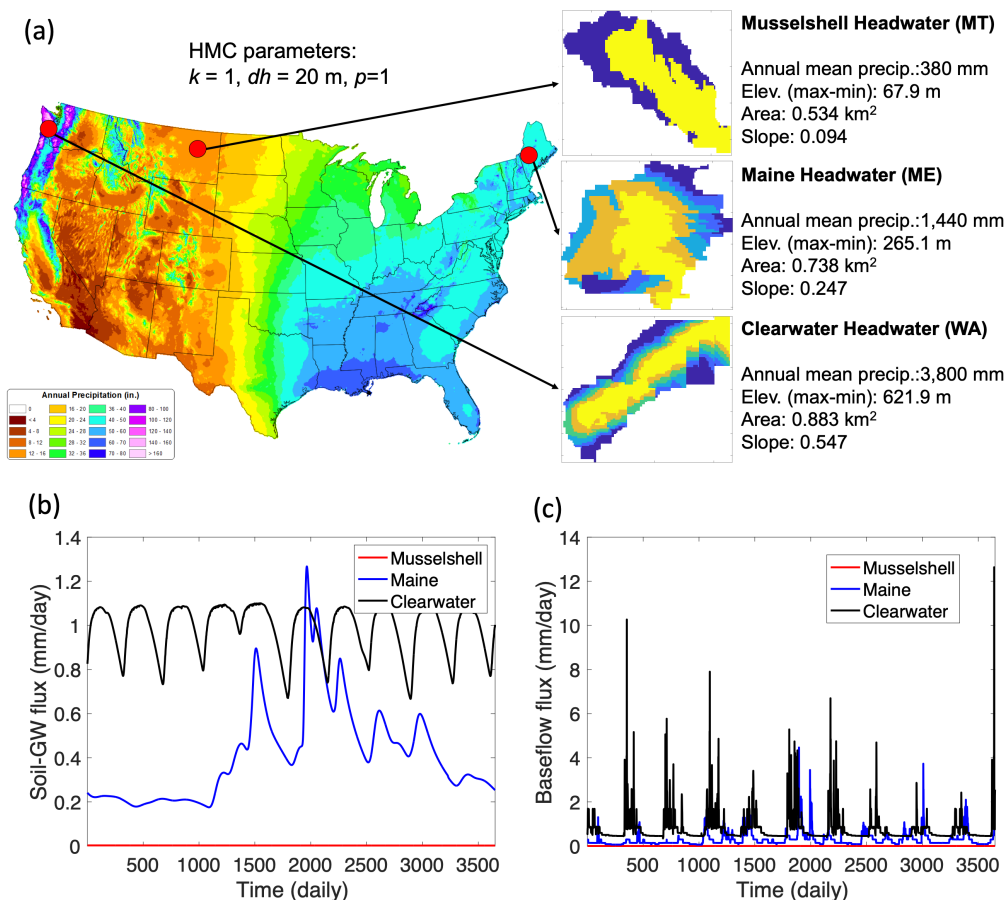
Figure 12. (a) the additional three sites are marked (red dots) on the PRISM 30-year normal precipitation map (2022, PRISM Group, Oregon State University). The sites include Musselshell headwater (MT), Maine headwater (ME), and Clearwater headwater (WA) catchments, (b) the LM4-SHARC's daily soil-groundwater 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

777   established by means of simple regression analysis or machine learning (ML) techniques (e.g., random forest)

778   (Zecharias and Brutsaert, 1988; Hong and Mohanty, 2023b; Tashie et al., 2021; Cai et al., 2021). ML procedure can

779   leverage remotely sensed hydrologic data, such as surface soil moisture (e.g., SMAP), evapotranspiration (e.g.,

780   MODIS), and groundwater storage (e.g., GRACE), to enhance the robustness of groundwater parameter estimation

781   with identified relative importance of each input variable. Moreover, the recent launch of SWOT (Surface Water and

782   Ocean Topography) spaceborne mission offers the potential to gather river discharge and baseflow data at a

783   temporal resolution of interest (e.g., daily) with unprecedented global coverage (Baratelli et al., 2018; Li et al., 2020;

784   Wongchuig-Correa et al., 2020). These studies aimed to overcome the spatial and temporal gaps in SWOT

785   observations by integrating a large-scale hydrologic model with synthetic SWOT data through techniques like data

786   assimilation. Based on existing findings about the utility of SWOT data for the baseflow estimation, the possibility

787   of obtaining baseflow and streamflow at a higher spatial and temporal resolution will be further explored to use

788   model-derived outputs as surrogate data to investigate the catchment-scale recession behavior with global coverage.

789

790   **5.   Conclusions**

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

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

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

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

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

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

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

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

799   observationally-derived stream discharge information.

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

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

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

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

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

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

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

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

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

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

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

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

812   significantly influences the dynamics of land surface energy balance.

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

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

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

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

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

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

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

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

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

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

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

824   as flooding with enhanced streamflow production.

825

Author contributions

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

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

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

830

Competing interests

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

833

840

Code and data availability

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

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

844

References

846   Andreae, M. O., Artaxo, P., Brandao, C., Carswell, F. E., Ciccioli, P., Da Costa, A., Culf, A. D., Esteves, J., Gash, J.
847   H., and Grace, J.: Biogeochemical cycling of carbon, water, energy, trace gases, and aerosols in Amazonia:
848   The LBA-EUSTACH experiments, J. Geophys. Res.: Atmos., 107, LBA 33-31-LBA 33-25, 2002.
849   Bales, R., Stacy, E., Safeeq, M., Meng, X., Meadows, M., Oroza, C., Conklin, M., Glaser, S., and Wagenbrenner, J.:
850   Spatially distributed water-balance and meteorological data from the rain–snow transition, southern Sierra
851   Nevada, California, Earth Syst. Sci. Data, 10, 1795-1805, 10.5194/essd-10-1795-2018, 2018.

852 Baratelli, F., Flipo, N., Rivière, A., and Biancamaria, S.: Retrieving river baseflow from SWOT spaceborne mission,
853   Remote Sens. Environ., 218, 44-54, 2018.
854 Baroni, G., Schalge, B., Rakovec, O., Kumar, R., Schüler, L., Samaniego, L., Simmer, C., and Attinger, S.: A
855   Comprehensive Distributed Hydrological Modeling Intercomparison to Support Process Representation and
856   Data Collection Strategies, Water Resour. Res., 55, 990-1010, 10.1029/2018wr023941, 2019.
857 Basha, H. A.: Traveling wave solution of the Boussinesq equation for groundwater flow in horizontal aquifers,
858   Water Resour. Res., 49, 1668-1679, 10.1002/wrcr.20168, 2013.
859 Best, M. J., Pryor, M., Clark, D., Rooney, G. G., Essery, R., Ménard, C., Edwards, J., Hendry, M., Porson, A., and
860   Gedney, N.: The Joint UK Land Environment Simulator (JULES), model description–Part 1: energy and
861   water fluxes, Geosci. Model Dev., 4, 677-699, 2011.
862 Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology/Un modèle
863   à base physique de zone d'appel variable de l'hydrologie du bassin versant, Hydrol. Sci. J., 24, 43-69, 1979.
864 Bisht, G., Huang, M. Y., Zhou, T., Chen, X. Y., Dai, H., Hammond, G. E., Riley, W. J., Downs, J. L., Liu, Y., and
865   Zachara, J. M.: Coupling a three-dimensional subsurface flow and transport model with a land surface
866   model to simulate stream-aquifer-land interactions (CP v1.0), Geosci. Model Dev., 10, 4539-4562,
867   10.5194/gmd-10-4539-2017, 2017.
868 Blyth, E. M., Arora, V. K., Clark, D. B., Dadson, S. J., De Kauwe, M. G., Lawrence, D. M., Melton, J. R., Pongratz,
869   J., Turton, R. H., Yoshimura, K., and Yuan, H.: Advances in Land Surface Modelling, Curr. Clim. Change
870   Rep., 7, 45-71, 10.1007/s40641-021-00171-5, 2021.
871 Boussinesq, J.: Sur le de´bit, en temps de se´cheresse, d'une source alimente´e par une nappe d'eaux d'infiltration,
872   C.R. Hebd Se´-ances Acad. Sci., 136, 1511-1517, 1903.
873 Brutsaert, W.: The unit response of groundwater outflow from a hillslope, Water Resour. Res., 30, 2759-2763,
874   10.1029/94wr01396, 1994.
875 Brutsaert, W. and Lopez, J. P.: Basin-scale geohydrologic drought flow features of riparian aquifers in the Southern
876   Great Plains, Water Resour. Res., 34, 233-240, 10.1029/97wr03068, 1998.
877 Brutsaert, W. and Nieber, J. L.: Regionalized drought flow hydrographs from a mature glaciated plateau, Water
878   Resour. Res., 13, 637-643, 10.1029/wr013i003p00637, 1977.
879 Cai, H., Shi, H., Liu, S., and Babovic, V.: Impacts of regional characteristics on improving the accuracy of
880   groundwater level prediction using machine learning: The case of central eastern continental United States,
881   J. Hydrol.: Reg. Stud., 37, 100930, 2021.
882 Campoy, A., Ducharne, A., Cheruy, F., Hourdin, F., Polcher, J., and Dupont, J.: Response of land surface fluxes and
883   precipitation to different soil bottom hydrological conditions in a general circulation model, J. Geophys.
884   Res.: Atmos., 118, 10,725-710,739, 2013.
885 Chaney, N. W., Metcalfe, P., and Wood, E. F.: HydroBlocks: a field-scale resolving land surface model for
886   application over continental extents, Hydrol. Processes, 30, 3543-3559, 10.1002/hyp.10891, 2016.
887 Chaney, N. W., Torres-Rojas, L., Vergopolan, N., and Fisher, C. K.: HydroBlocks v0.2: enabling a field-scale two-
888   way coupling between the land surface and river networks in Earth system models, Geosci. Model Dev.,
889   14, 6813-6832, 10.5194/gmd-14-6813-2021, 2021.
890 Chaney, N. W., Van Huijgevoort, M. H. J., Shevliakova, E., Malyshev, S., Milly, P. C. D., Gauthier, P. P. G., and
891   Sulman, B. N.: Harnessing big data to rethink land heterogeneity in Earth system models, Hydrol. Earth
892   Syst. Sci., 22, 3311-3330, 10.5194/hess-22-3311-2018, 2018.
893 Chen, X. and Hu, Q.: Groundwater influences on soil moisture and surface evaporation, J. Hydrol., 297, 285-300,
894   2004.
895 Cheng, S., Cheng, L., Liu, P., Qin, S., Zhang, L., Xu, C. Y., Xiong, L., Liu, L., and Xia, J.: An Analytical Baseflow
896   Coefficient Curve for Depicting the Spatial Variability of Mean Annual Catchment Baseflow, Water
897   Resour. Res., 57, 10.1029/2020wr029529, 2021.
898 Clark, M. P., Rupp, D. E., Woods, R. A., Tromp-Van Meerveld, H. J., Peters, N. E., and Freer, J. E.: Consistency
899   between hydrological models and field observations: linking processes at the hillslope scale to hydrological
900   responses at the watershed scale, Hydrol. Processes, 23, 311-319, 10.1002/hyp.7154, 2009.
901 Clark, M. P., Fan, Y., Lawrence, D. M., Adam, J. C., Bolster, D., Gochis, D. J., Hooper, R. P., Kumar, M., Leung,
902   L. R., Mackay, D. S., Maxwell, R. M., Shen, C. P., Swenson, S. C., and Zeng, X. B.: Improving the
903   representation of hydrologic processes in Earth System Models, Water Resour. Res., 51, 5929-5956,
904   10.1002/2015wr017096, 2015.
905 De Rosnay, P., Polcher, J. d., Bruen, M., and Laval, K.: Impact of a physically based soil water flow and soil-plant
906   interaction representation for modeling large-scale land surface processes, J. Geophys. Res.: Atmos., 107,
907   ACL 3-1-ACL 3-19, 2002.

908   De St Venant, B.: Theorie du mouvement non-permanent des eaux avec application aux crues des rivers et a
909   l'introduntion des Marees dans leur lit, Academic de Sci. Comptes Redus, 73, 148-154, 1871.
910   Delworth, T. L., Cooke, W. F., Adcroft, A., Bushuk, M., Chen, J. H., Dunne, K. A., Ginoux, P., Gudgel, R.,
911   Hallberg, R. W., and Harris, L.: SPEAR: The next generation GFDL modeling system for seasonal to
912   multidecadal prediction and projection, J. Adv. Model. Earth Syst., 12, e2019MS001895, 2020.
913   Dunne, J. P., Horowitz, L., Adcroft, A., Ginoux, P., Held, I., John, J., Krasting, J. P., Malyshev, S., Naik, V., and
914   Paulot, F.: The GFDL Earth System Model version 4.1 (GFDL-ESM 4.1): Overall coupled model
915   description and simulation characteristics, J. Adv. Model. Earth Syst., 12, e2019MS002015, 2020.
916   Dupuit, J.: Etudes theoriques et pratiques sur le mouvement des eaux dans les canaux decouverts et a travers les
917   terrains permeables avec des considerations relatives au regime des grandes eaux, au debouche a leur
918   donner et a la marche des alluvions dans les rivieres a fond mobile, Dunod1863.
919   Fan, Y.: Groundwater in the Earth's critical zone: Relevance to large-scale patterns and processes, Water Resour.
920   Res., 51, 3052-3069, 10.1002/2015wr017037, 2015.
921   Fan, Y. and Miguez-Macho, G.: A simple hydrologic framework for simulating wetlands in climate and earth
922   system models, Clim. Dyn., 37, 253-278, 2011.
923   Fan, Y., Miguez-Macho, G., Weaver, C. P., Walko, R., and Robock, A.: Incorporating water table dynamics in
924   climate modeling: 1. Water table observations and equilibrium water table simulations, Journal of
925   Geophysical Research-Atmospheres, 112, 10.1029/2006jd008111, 2007.
926   Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., Brooks, P. D., Dietrich, W. E.,
927   Flores, A., Grant, G., Kirchner, J. W., Mackay, D. S., McDonnell, J. J., Milly, P. C. D., Sullivan, P. L.,
928   Tague, C., Ajami, H., Chaney, N., Hartmann, A., Hazenberg, P., McNamara, J., Pelletier, J., Perket, J.,
929   Rouholahnejad-Freund, E., Wagener, T., Zeng, X., Beighley, E., Buzan, J., Huang, M., Livneh, B.,
930   Mohanty, B. P., Nijssen, B., Safeeq, M., Shen, C., Verseveld, W., Volk, J., and Yamazaki, D.: Hillslope
931   Hydrology in Global Change Research and Earth System Modeling, Water Resour. Res., 55, 1737-1772,
932   10.1029/2018wr023903, 2019.
933   Forchheimer, P.: Uber die ergiebigkeit von brummen, Anlagen und Sickerschlitzen, Zeitsch Archit. Ing. Ver.,
934   Hannover, 32, 539-563, 1986.
935   Gentine, P., Green, J. K., Guérin, M., Humphrey, V., Seneviratne, S. I., Zhang, Y., and Zhou, S.: Coupling between
936   the terrestrial carbon and water cycles—a review, Environ. Res. Lett., 14, 083003, 2019.
937   Gleeson, T., Wang-Erlandsson, L., Porkka, M., Zipper, S. C., Jaramillo, F., Gerten, D., Fetzer, I., Cornell, S. E.,
938   Piemontese, L., Gordon, L. J., Rockström, J., Oki, T., Sivapalan, M., Wada, Y., Brauman, K. A., Flörke,
939   M., Bierkens, M. F. P., Lehner, B., Keys, P., Kummu, M., Wagener, T., Dadson, S., Troy, T. J., Steffen,
940   W., Falkenmark, M., and Famiglietti, J. S.: Illuminating water cycle modifications and Earth system
941   resilience in the Anthropocene, Water Resour. Res., 56, 10.1029/2019wr024957, 2020.
942   Gochis, D. J., M. Barlage, A. Dugger, K. FitzGerald, L. Karsten, M. McAllister, J. McCreight, J. Mills, A.
943   RafieeiNasab, L. Read, K. Sampson, D. Yates, W. Yu,: The NCAR WRF-Hydro Modeling System
944   Technical Description (Version 5.2), National Center for Atmospheric Research, 2021.
945   Gómez-Hernández, J. J. and Gorelick, S. M.: Effective groundwater model parameter values: Influence of spatial
946   variability of hydraulic conductivity, leakance, and recharge, Water Resour. Res., 25, 405-419, 1989.
947   Heath, R. C.: Basic ground-water hydrology, US Geological Survey2004.
948   Hong, M. and Mohanty, B.: A new method for effective parameterization of catchment-scale aquifer through event-
949   scale recession analysis, Adv. Water Resour., 174, 104408, 2023a.
950   Hong, M. and Mohanty, B. P.: Representing Bidirectional Hydraulic Continuum Between the Stream and Hillslope
951   in the National Water Model for Improved Streamflow Prediction, J. Adv. Model. Earth Syst., 15,
952   10.1029/2022ms003325, 2023b.
953   Hornberger, G. M. and Remson, I.: A MOVING BOUNDARY MODEL OF A ONE-DIMENSIONAL
954   SATURATED-UNSATURATED, TRANSIENT POROUS FLOW SYSTEM, Water Resour. Res., 6, 898-
955   +, 10.1029/WR006i003p00898, 1970.
956   Hunsaker, C. T., Whitaker, T. W., and Bales, R. C.: Snowmelt Runoff and Water Yield Along Elevation and
957   Temperature Gradients in California's Southern Sierra Nevada1, JAWRA J. Am. Water Resour. Assoc., 48,
958   667-678, 10.1111/j.1752-1688.2012.00641.x, 2012.
959   Jachens, E. R., Roques, C., Rupp, D. E., and Selker, J. S.: Recession analysis revisited: impacts of climate on
960   parameter estimation, Hydrol. Earth Syst. Sci., 24, 1159-1170, 10.5194/hess-24-1159-2020, 2020.
961   Jepsen, S. M., Harmon, T. C., Meadows, M. W., and Hunsaker, C. T.: Hydrogeologic influence on changes in
962   snowmelt runoff with climate warming: Numerical experiments on a mid-elevation catchment in the Sierra
963   Nevada, USA, J. Hydrol., 533, 332-342, 2016.

964   Jing, M., Heße, F., Kumar, R., Kolditz, O., Kalbacher, T., and Attinger, S.: Influence of input and parameter
965   uncertainty on the prediction of catchment-scale groundwater travel time distributions, Hydrol. Earth Syst.
966   Sci., 23, 171-190, 2019.
967   Karlsen, R. H., Bishop, K., Grabs, T., Ottosson-Löfvenius, M., Laudon, H., and Seibert, J.: The role of landscape
968   properties, storage and evapotranspiration on variability in streamflow recessions in a boreal catchment, J.
969   Hydrol., 570, 315-328, 10.1016/j.jhydrol.2018.12.065, 2019.
970   Kirchner, J. W.: Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling,
971   and doing hydrology backward, Water Resour. Res., 45, n/a-n/a, 10.1029/2008wr006912, 2009.
972   Kollet, S. J. and Maxwell, R. M.: Integrated surface-groundwater flow modeling: A free-surface overland flow
973   boundary condition in a parallel groundwater flow model, Adv. Water Resour., 29, 945-958,
974   10.1016/j.advwatres.2005.08.006, 2006.
975   Kollet, S. J. and Maxwell, R. M.: Capturing the influence of groundwater dynamics on land surface processes using
976   an integrated, distributed watershed model, Water Resour. Res., 44, 2008.
977   Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang,
978   Z. L., Levis, S., and Sakaguchi, K.: Parameterization improvements and functional and structural advances
979   in version 4 of the Community Land Model, J. Adv. Model. Earth Syst., 3, 2011.
980   Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B.,
981   Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J.,
982   Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., Broeke,
983   M., Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B.,
984   Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J.,
985   Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto,
986   D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Val Martin, M., and Zeng, X.:
987   The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of
988   Forcing Uncertainty, J. Adv. Model. Earth Syst., 11, 4245-4287, 10.1029/2018ms001583, 2019.
989   Leung, L. R., Huang, M., Qian, Y., and Liang, X.: Climate–soil–vegetation control on groundwater table dynamics
990   and its feedbacks in a climate model, Clim. Dyn., 36, 57-81, 2011.
991   Li, D., Andreadis, K. M., Margulis, S. A., and Lettenmaier, D. P.: A Data Assimilation Framework for Generating
992   Space-Time Continuous Daily SWOT River Discharge Data Products, Water Resour. Res., 56,
993   10.1029/2019wr026999, 2020.
994   Li, H., Wigmosta, M. S., Wu, H., Huang, M., Ke, Y., Coleman, A. M., and Leung, L. R.: A physically based runoff
995   routing model for land surface and earth system models, J. Hydrometeorol., 14, 808-828, 2013.
996   Liang, X., Xie, Z. H., and Huang, M. Y.: A new parameterization for surface and groundwater interactions and its
997   impact on water budgets with the variable infiltration capacity (VIC) land surface model, J. Geophys. Res.:
998   Atmos., 108, 10.1029/2002jd003090, 2003.
999   Manning, R., Griffith, J. P., Pigot, T., and Vernon-Harcourt, L. F.: On the flow of water in open channels and pipes,
1000   1890.
1001 Maxwell, R. M. and Kollet, S. J.: Interdependence of groundwater dynamics and land-energy feedbacks under
1002   climate change, Nat. Geosci., 1, 665-669, 10.1038/ngeo315, 2008.
1003 Maxwell, R. M., Chow, F. K., and Kollet, S. J.: The groundwater–land-surface–atmosphere connection: Soil
1004   moisture effects on the atmospheric boundary layer in fully-coupled simulations, Adv. Water Resour., 30,
1005   2447-2466, 10.1016/j.advwatres.2007.05.018, 2007.
1006 Maxwell, R. M., Condon, L. E., and Kollet, S. J.: A high-resolution simulation of groundwater and surface water
1007   over most of the continental US with the integrated hydrologic model ParFlow v3, Geosci. Model Dev., 8,
1008   923-937, 10.5194/gmd-8-923-2015, 2015.
1009 Maxwell, R. M., Lundquist, J. K., Mirocha, J. D., Smith, S. G., Woodward, C. S., and Tompson, A. F. B.:
1010   Development of a Coupled Groundwater–Atmosphere Model, Mon. Weather Rev., 139, 96-116,
1011   10.1175/2010mwr3392.1, 2011.
1012 McCabe, M., Wood, E. F., Wójcik, R., Pan, M., Sheffield, J., Gao, H., and Su, H.: Hydrological consistency using
1013   multi-sensor remote sensing data for water and energy cycle studies, Remote Sens. Environ., 112, 430-444,
1014   2008.
1015 Miguez-Macho, G., Fan, Y., Weaver, C. P., Walko, R., and Robock, A.: Incorporating water table dynamics in
1016   climate modeling: 2. Formulation, validation, and soil moisture simulation, J. Geophys. Res.: Atmos., 112,
1017   10.1029/2006jd008112, 2007.
1018 Miguez-Macho, G. and Fan, Y.: The role of groundwater in the Amazon water cycle: 1. Influence on seasonal
1019   streamflow, flooding and wetlands, J. Geophys. Res.: Atmos., 117, n/a-n/a, 10.1029/2012jd017539, 2012.

1020 Milly, P. C., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., and Stouffer, R.
1021   J.: Stationarity is dead: Whither water management?, Science, 319, 573-574, 2008.
1022 Milly, P. C. D., Malyshev, S. L., Shevliakova, E., Dunne, K. A., Findell, K. L., Gleeson, T., Liang, Z., Phillipps, P.,
1023   Stouffer, R. J., and Swenson, S.: An Enhanced Model of Land Water and Energy for Global Hydrologic
1024   and Earth-System Studies, J. Hydrometeorol., 15, 1739-1761, 10.1175/jhm-d-13-0162.1, 2014.
1025 Mu, Q., Zhao, M., and Running, S. W.: Evolution of hydrological and carbon cycles under a changing climate,
1026   Hydrol. Processes, 25, 4093-4102, 10.1002/hyp.8367, 2011.
1027 Newman, A. J., Clark, M. P., Winstral, A., Marks, D., and Seyfried, M.: The Use of Similarity Concepts to
1028   Represent Subgrid Variability in Land Surface Models: Case Study in a Snowmelt-Dominated Watershed,
1029   J. Hydrometeorol., 15, 1717-1738, 10.1175/jhm-d-13-038.1, 2014.
1030 Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D.,
1031   Rosero, E., Tewari, M., and Xia, Y.: The community Noah land surface model with multiparameterization
1032   options (Noah-MP): 1. Model description and evaluation with local-scale measurements, J. Geophys. Res.,
1033   116, 10.1029/2010jd015139, 2011.
1034 Oleson, K. W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., Levis, S., Li, F., Riley, W.
1035   J., Subin, Z. M., Swenson, S. C., and Thornton, P. E.: Technical Description of version 4.5 of the
1036   Community Land Model (CLM), NCAR Earth System Laboratory, 10.5065/D6RR1W7M, 2013.
1037 Ouyang, L., Liu, S., Ye, J., Liu, Z., Sheng, F., Wang, R., and Lu, Z.: Quantitative assessment of surface runoff and
1038   base flow response to multiple factors in Pengchongjian small watershed, Forests, 9, 553, 2018.
1039 Pappenberger, F., Dutra, E., Wetterhall, F., and Cloke, H. L.: Deriving global flood hazard maps of fluvial floods
1040   through a physical model cascade, Hydrol. Earth Syst. Sci., 16, 4143-4156, 2012.
1041 Pauritsch, M., Birk, S., Wagner, T., Hergarten, S., and Winkler, G.: Analytical approximations of discharge
1042   recessions for steeply sloping aquifers in alpine catchments, Water Resour. Res., 51, 8729-8740,
1043   10.1002/2015wr017749, 2015.
1044 Pokhrel, Y. N., Fan, Y., Miguez-Macho, G., Yeh, P. J. F., and Han, S.-C.: The role of groundwater in the Amazon
1045   water cycle: 3. Influence on terrestrial water storage computations and comparison with GRACE, J.
1046   Geophys. Res.: Atmos., 118, 3233-3244, 10.1002/jgrd.50335, 2013.
1047 Rupp, D. E. and Selker, J. S.: On the use of the Boussinesq equation for interpreting recession hydrographs from
1048   sloping aquifers, Water Resour. Res., 42, n/a-n/a, 10.1029/2006wr005080, 2006.
1049 Safeeq, M. and Hunsaker, C. T.: Characterizing Runoff and Water Yield for Headwater Catchments in the Southern
1050   Sierra Nevada, JAWRA J. Am. Water Resour. Assoc., 52, 1327-1346, 10.1111/1752-1688.12457, 2016.
1051 Shen, C. P. and Phanikumar, M. S.: A process-based, distributed hydrologic model based on a large-scale method
1052   for surface-subsurface coupling, Adv. Water Resour., 33, 1524-1541, 10.1016/j.advwatres.2010.09.002,
1053   2010.
1054 Sivapalan, M.: Pattern, process and function: elements of a unified theory of hydrology at the catchment scale,
1055   Encyclopedia of hydrological sciences, 2006.
1056 Strack, O. D. L.: A Dupuit-Forchheimer Model for three-dimensional flow with variable density, Water Resour.
1057   Res., 31, 3007-3017, 10.1029/95wr02254, 1995.
1058 Strelkoff, T.: Numerical solution of Saint-Venant equations, J. Hydraul. Div., 96, 223-252, 1970.
1059 Subin, Z. M., Milly, P. C. D., Sulman, B. N., Malyshev, S., and Shevliakova, E.: Resolving terrestrial ecosystem
1060   processes along a subgrid topographic gradient for an earth-system model, 10.5194/hessd-11-8443-2014,
1061   2014.
1062 Szilagyi, J. and Parlange, M. B.: Baseflow separation based on analytical solutions of the Boussinesq equation, J.
1063   Hydrol., 204, 251-260, 10.1016/s0022-1694(97)00132-7, 1998.
1064 Szilagyi, J., Parlange, M. B., and Albertson, J. D.: Recession flow analysis for aquifer parameter determination,
1065   Water Resour. Res., 34, 1851-1857, 10.1029/98wr01009, 1998.
1066 Takata, K., Emori, S., and Watanabe, T.: Development of the minimal advanced treatments of surface interaction
1067   and runoff, Global Planet. Change, 38, 209-222, 2003.
1068 Tashie, A., Pavelsky, T., and Band, L. E.: An Empirical Reevaluation of Streamflow Recession Analysis at the
1069   Continental Scale, Water Resour. Res., 56, 10.1029/2019wr025448, 2020.
1070 Tashie, A., Pavelsky, T., Band, L., and Topp, S.: Watershed-Scale Effective Hydraulic Properties of the Continental
1071   United States, J. Adv. Model. Earth Syst., 13, 10.1029/2020ms002440, 2021.
1072 Troch, P. A., De Troch, F. P., and Brutsaert, W.: Effective water table depth to describe initial conditions prior to
1073   storm rainfall in humid regions, Water Resour. Res., 29, 427-434, 10.1029/92wr02087, 1993.
1074 Troch, P. A., Berne, A., Bogaart, P., Harman, C., Hilberts, A. G. J., Lyon, S. W., Paniconi, C., Pauwels, V. R. N.,
1075   Rupp, D. E., Selker, J. S., Teuling, A. J., Uijlenhoet, R., and Verhoest, N. E. C.: The importance of

1076   hydraulic groundwater theory in catchment hydrology: The legacy of Wilfried Brutsaert and Jean-Yves
1077   Parlange, Water Resour. Res., 49, 5099-5116, 10.1002/wrcr.20407, 2013.
1078 USGS: 3D Elevation Program 1-Meter Resolution Digital Elevation Model [dataset], 2019.
1079 Vannier, O., Braud, I., and Anquetin, S.: Regional estimation of catchment-scale soil properties by means of
1080   streamflow recession analysis for use in distributed hydrological models, Hydrol. Processes, 28, 6276-
1081   6291, 10.1002/hyp.10101, 2014.
1082 Wongchuig-Correa, S., de Paiva, R. C. D., Biancamaria, S., and Collischonn, W.: Assimilation of future SWOT-
1083   based river elevations, surface extent observations and discharge estimations into uncertain global
1084   hydrological models, J. Hydrol., 590, 125473, 2020.
1085 Xu, B., Ye, M., Dong, S., Dai, Z., and Pei, Y.: A new model for simulating spring discharge recession and
1086   estimating effective porosity of karst aquifers, J. Hydrol., 562, 609-622, 2018.
1087 Xu, Z., Siirila-Woodburn, E. R., Rhoades, A. M., and Feldman, D.: Sensitivities of subgrid-scale physics schemes,
1088   meteorological forcing, and topographic radiation in atmosphere-through-bedrock integrated process
1089   models: a case study in the Upper Colorado River basin, Hydrol. Earth Syst. Sci., 27, 1771-1789, 2023.
1090 Yeh, P. J. F. and Eltahir, E. A. B.: Representation of Water Table Dynamics in a Land Surface Scheme. Part I:
1091   Model Development, Journal of Climate, 18, 1861-1880, 10.1175/jcli3330.1, 2005.
1092 Zecharias, Y. B. and Brutsaert, W.: Recession characteristics of groundwater outflow and base flow from
1093   mountainous watersheds, Water Resour. Res., 24, 1651-1658, 1988.
1094 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
1095   of stream-aquifer water interaction: a case study of the Heihe River basin, northwestern China, Hydrol.
1096   Earth Syst. Sci., 20, 2333-2352, 10.5194/hess-20-2333-2016, 2016.
1097 Zhao, M., Golaz, J. C., Held, I., Guo, H., Balaji, V., Benson, R., Chen, J. H., Chen, X., Donner, L., and Dunne, J.:
1098   The GFDL global atmosphere and land model AM4. 0/LM4. 0: 2. Model description, sensitivity studies,
1099   and tuning strategies, J. Adv. Model. Earth Syst., 10, 735-769, 2018.
1100 Zorzetto, E., Malyshev, S., Chaney, N., Paynter, D., Menzel, R., and Shevliakova, E.: Effects of complex terrain on
1101   the shortwave radiative balance: a sub-grid-scale parameterization for the GFDL Earth System Model
1102   version 4.1, Geosci. Model Dev., 16, 1937-1960, 10.5194/gmd-16-1937-2023, 2023.
1103