

# Water Resources Research

## RESEARCH ARTICLE

10.1029/2020WR027714



### Key Points:

- We compare the simulated hydrology of a semi-idealized crystalline watershed using six common conceptual models of the mountain subsurface
- Discharge and flow path distributions are most sensitive to the choice of properties in the shallow and deeper subsurface, respectively
- Discharge, water tables, and flow paths respond differently to conceptual models and model evaluation on one metric risks biasing outputs

### Correspondence to:

G. A. Rapp,  
grapp@email.arizona.edu

### Citation:

Rapp, G. A., Condon, L. E., & Markovich, K. H. (2020). Sensitivity of simulated mountain block hydrology to subsurface conceptualization. *Water Resources Research*, 56, e2020WR027714. <https://doi.org/10.1029/2020WR027714>

Received 12 APR 2020

Accepted 29 SEP 2020

Accepted article online 3 OCT 2020

## Sensitivity of Simulated Mountain Block Hydrology to Subsurface Conceptualization

G. A. Rapp<sup>1</sup> , L. E. Condon<sup>1</sup> , and K. H. Markovich<sup>1</sup> 

<sup>1</sup>Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA

**Abstract** Mountain block systems are critical to water resources and have been heavily studied and modeled in recent decades. However, due to lack of field data, there is little consistency in how models represent the mountain block subsurface. While there is a large body of research on subsurface heterogeneity, few studies have evaluated the effect that common conceptual choices modelers make in mountainous systems have on simulated hydrology. Here we simulate the hydrology of a semi-idealized headwater catchment using six common conceptual models of the mountain block subsurface. These scenarios include multiple representations of hydraulic conductivity decaying with depth, changes in soil depth with topography, and anisotropy. We evaluate flow paths, discharge, and water tables to quantify the impact of subsurface conceptualization on hydrologic behavior in three dimensions. Our results show that adding higher conductivity layers in the shallow subsurface concentrates flow paths near the surface and increases average saturated flow path velocities. Increasing heterogeneity by adding additional layers or introducing anisotropy increases the variance in the relationship between the age and length of saturated flow paths. Discharge behavior is most sensitive to heterogeneity in the shallow subsurface layers. Water tables are less sensitive to layering than they are to the overall conductivity in the domain. Anisotropy restricts flow path depths and controls discharge from storage but has little effect on governing runoff. Differences in the response of discharge, water table depth, and residence time distribution to subsurface representation highlight the need to consider model applications when determining the level of complexity that is needed.

**Plain Language Summary** Computer models are a commonly used tool to understand and predict the behavior of mountainous watersheds. However, modeling groundwater is uniquely challenging in these systems because groundwater flow is strongly influenced by the geometry of the bedrock, yet observations of depth to bedrock and measurements of hydraulic properties with depth are infrequent. Modelers commonly rely on simplifying assumptions to represent the geologic layering in mountainous watersheds. However, the impact of these assumptions on simulated hydrology is rarely evaluated. Our study compares six commonly used representations of the subsurface to explore how basic choices in model construction influence the resulting hydrology. The six models vary in complexity from no layering at all to more complex variations in soil depth. We compare the behavior of our six cases based on groundwater levels, streamflow, and groundwater flow pathways. We show that streamflow changes the most due to changes in the model layers closest to the surface. The speed of groundwater flow pathways is more sensitive to the properties of layers at all depths. Differences in how streamflow and groundwater flow pathways respond to various representations highlight the fact that good model performance with respect to one variable can hide uncertainty in other areas.

## 1. Introduction

Mountain regions are important water sources for much of the world, and understanding headwater hydrology is vital to anticipating the effects of human development and climate change on these resources (Immerzeel et al., 2019). However, the geologic complexity of mountain systems makes them unique from other types of watersheds and limits our ability to translate lessons learned in other settings to mountainous watersheds. Thinly layered soils, steep topography, and complex folded and faulted geology affect the runoff and infiltration processes in the mountain block, driving the storage-discharge behavior of these systems in unique ways (Ajami et al., 2011).

Our current conceptual models of mountain hydrogeology are limited by available observations. Field measurements are difficult to gather at a sufficient spatial and temporal resolution to support watershed-scale

©2020. The Authors

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

subsurface characterization. Drilling into the mountain block to gather subsurface data can be expensive, and the hydrologic properties that hydrogeologists must infer and extrapolate from geologic observations can be uncertain. Given these limitations, computer models are commonly used tools to explore mountain groundwater hydrology (Markovich et al., 2019). Still, to build these models, the modelers must make assumptions about the hydrogeology of the mountain block. While modelers routinely calibrate or explore the hydrologic sensitivity to the parameter values they assign to their hydrogeologic units (e.g., porosity and hydraulic conductivity [ $K$ ]), the more conceptual assumptions modelers make when developing the subsurface configuration (e.g., layer thicknesses) of headwater hydrologic models have not been well explored.

The current understanding of mountain groundwater systems builds from early work on regional groundwater systems. Some of the first models of regional groundwater systems explored the effects of topography on groundwater table configuration, assuming the form of the water table in advance (Freeze & Witherspoon, 1967; Tóth, 1963). Tóth (1963) demonstrated the existence of nested flow paths within a regional groundwater system, which he categorized as local, intermediate, and regional flow paths. His numerical experiments involved many assumptions, including shallow flank slopes, negligible longitudinal flow (justifying the use of a two-dimensional domain), an isotropic, heterogeneous subsurface, and, most significantly, that the water table was a subdued replica of topography. Despite the highly simplified nature of his theoretical analysis, Tóth's nested flow path model remains the foundation for many of our groundwater conceptualizations today.

Building upon Tóth's seminal work, hydrogeologists have since explored the drivers of regional groundwater flow. Haitjema and Mitchell-Bruker (2005) developed a two-dimensional model to evaluate the underlying assumptions that lead to Tóth's model of the water table as a subdued replica of the topography. They relaxed many of Tóth's assumptions, including the shape of the water table, the thickness of the active aquifer, and the variability of the topography. They systematically varied levels of recharge and  $K$  across their domain and demonstrated that the water table configuration is primarily controlled by the relationship between the average recharge and  $K$ . A relatively low ratio of average recharge to  $K$  indicates a recharge-controlled water table, which is deeper and results in more prominent horizontal flow and regional flow paths. Conversely, a relatively high ratio of average recharge to  $K$  indicates a topography-controlled water table, where local and intermediate flow paths are more prominent. Haitjema and Mitchell-Bruker (2005) demonstrated that the water table configuration used by Tóth is not universal but representative of a system with a high ratio of recharge to  $K$ .

While the work of Haitjema and Mitchell-Bruker (2005) was an important advance from Tóth's analysis, their work was still based on two-dimensional homogenous domains. Two-dimensional modeling of regional groundwater systems can lead to fundamentally different water tables and flow paths than what occurs in three dimensions (J.-Z. Wang et al., 2017). In addition, mountain block systems are geologically complex; faulting, fracturing, and differential weathering can lead to sharp gradients in  $K$ —potentially spanning several orders of magnitude in the shallow (<100 m) subsurface alone (Welch & Allen, 2014). Furthermore, steep topography can increase the relative importance of longitudinal flow, or flow that is perpendicular to the regional flow direction, resulting in more complex nested flow paths in three dimensions (Gleeson & Manning, 2008). This combination of subsurface heterogeneity and steep topography makes common simplifying assumptions about groundwater flow more difficult to justify. For example, the Dupuit-Forchheimer approximation assumes the resistance to vertical flow; we often lack the field data in the mountain block to determine its validity when calculating flow behavior.

Several studies have explored the controls of groundwater behavior specifically in mountain watershed systems. Forster and Smith (1988) conducted sensitivity analyses on mountain hydrology using a two-dimensional model to test topographic, geologic, climatic, and thermal effects on mountain groundwater systems. They determined that  $K$  is the most important driver for changes in mountain flow systems of the parameters they modeled. Gleeson and Manning (2008) expanded Tóth's regional groundwater flow analysis to a three-dimensional mountain block watershed. They concluded that Haitjema and Mitchell-Bruker's (2005) criteria for recharge versus topography-controlled water tables were generally valid in mountainous systems, but they further showed that relative importance of regional flow paths is modulated by the degree of incision in first-order streams. Welch and Allen (2012) demonstrated that the groundwater flow patterns (i.e., recharge zones and baseflow) that arise from topography-driven systems are generally

consistent across different configurations of topography. Ameli et al. (2018) corroborated these relationships between regional groundwater flow, recharge, and topographic relief. They combined field data and a semi-analytical flow and transport model showing that the proportion of regional flow paths increases with decreasing recharge.

While this body of research has advanced our understanding of mountain block hydrology, much of the work incorporates important assumptions. These include (1) simplifying systems to two dimensions (Forster & Smith, 1988; Welch & Allen, 2014), (2) assuming homogeneous domains or simple layering (Gleeson & Manning, 2008; Welch & Allen, 2012), (3) assuming steady-state solutions (Gleeson & Manning, 2008; C. Wang et al., 2018), and (4) ignoring or simplifying unsaturated flow (Ameli et al., 2018; Welch & Allen, 2014). These assumptions are valuable for conducting controlled numerical studies of basic interactions and have been critical in the advancement of our basic conceptualizations of mountainous systems. However, our increased ability to gather data (e.g., through geophysics and remote sensing) and the advances in numerical approaches that can leverage high-performance computing allow us to represent more complexity in the mountain block (Ball et al., 2014; Gilbert & Maxwell, 2017). We need to more thoroughly interrogate the connections between topography, subsurface complexity, and the simulated hydrologic behavior of the mountain block with fewer assumptions to help guide this work.

Currently, there is substantial variability in how modelers choose to represent mountain block watersheds and little research on the impact of the choice of representation. Here we present three of the most common approaches to representing mountain block watersheds in three-dimensional models. The simplest approach is to represent the mountain block watershed with two homogeneous layers to represent active (i.e., participating in the hydrologic system on human rather than geologic timescales) and inactive regions (Gleeson & Manning, 2008; Mayo et al., 2003; Welch & Allen, 2012). Building off of the first approach, some studies employ a combination of layers representing soil, saprolite, fractured bedrock, and unfractured bedrock, based on Welch and Allen's (2014) thorough review of crystalline mountain block hydrogeology (Gilbert & Maxwell, 2017; H. M. Voeckler et al., 2014). Alternatively, a common conceptual model with more complexity assumes an exponential decay of  $K$  with depth (Ameli et al., 2016; C. Wang et al., 2018), based on the empirical formula proposed by Louis (1972). In locations with more available soil and geologic data, some modelers add more complexity to these three conceptual models by varying layer thicknesses spatially (Engdahl & Maxwell, 2015; H. M. Voeckler et al., 2014).

There has been little evaluation of the effects of these various conceptual models on the modeled hydrologic behaviors to date. This study employs a model with transient, three-dimensional, variably saturated flow to identify systematic differences in watershed behaviors that result from common approaches to conceptualizing the subsurface. Specifically, we ask how sensitive is the simulated hydrologic behavior of the mountain block watershed to the common choices of conceptual models? To answer this, we use a semi-idealized mountainous domain to simulate scenarios of six common conceptual models of the mountain block. These include a scenario with a homogeneous subsurface, scenarios with several constant thickness layers, a scenario with variable soil depth, a scenario with anisotropy, and a scenario with exponential decay in  $K$  with depth.

We evaluate the resulting watershed behavior in each scenario with respect to water table depth, discharge patterns, and subsurface flow paths and discuss the sensitivities of these behaviors to the chosen conceptual model. Water tables and discharge patterns are commonly used metrics to evaluate a watershed model and are valuable for evaluating the storage and discharge behaviors of a watershed. Here we also analyze how the subsurface conceptual model influences the distribution of flow paths and groundwater ages. Groundwater flow paths can reveal sensitivities in watershed dynamics that may be obscured by aggregated water table or discharge response (Hale & McDonnell, 2016) and may be important for simulating nonstationary changes in watershed dynamics (Markovich et al., 2016; Tetzlaff et al., 2015).

## 2. Materials and Methods

This study uses an integrated hydrologic model coupled with particle tracking to simulate groundwater configuration, runoff, flow path lengths, and flow path ages and compares the impact of competing conceptual models on watershed simulations. All tests are applied in a semi-idealized mountain block domain. This section describes our modeling tools (section 2.1), the model domain (section 2.1.1), scenario designs (section 2.2), the overall simulation process (section 2.3), and model evaluation (section 2.4).

## 2.1. Modeling Tools

We use the integrated hydrologic model ParFlow-CLM to simulate the hydrologic and land surface processes of the mountain block (Ashby & Falgout, 1996; Jones & Woodward, 2001; Kollet & Maxwell, 2006, 2008; Maxwell, 2013; Maxwell & Miller, 2005). ParFlow simultaneously solves the governing equations for overland flow and groundwater flow in three dimensions, both in the saturated zone and the unsaturated zone. Overland flow is simulated with the kinematic wave approximation for each cell that has water ponded on the surface of the domain, and groundwater flow is simulated by solving the mixed form of Richards's equation. This work uses ParFlow coupled with a version of the Community Land Model (CLM) that simulates the land surface energy and water balance (Kollet & Maxwell, 2008). CLM uses hourly atmospheric variables including precipitation, solar radiation, wind speed, temperature, and specific humidity to simulate land surface processes and solve the water energy balance at the land surface.

To quantify the differences in flow paths between scenarios, we employ the EcoSLIM particle tracking code (Maxwell et al., 2019). EcoSLIM is a Lagrangian particle tracking model that uses gridded outputs from ParFlow-CLM (three-dimensional subsurface flow velocity, pressure, saturation, precipitation, and evapotranspiration) to insert, remove, and advect particles through the subsurface at each time step. EcoSLIM includes both advective and diffusive particle transport through the subsurface.

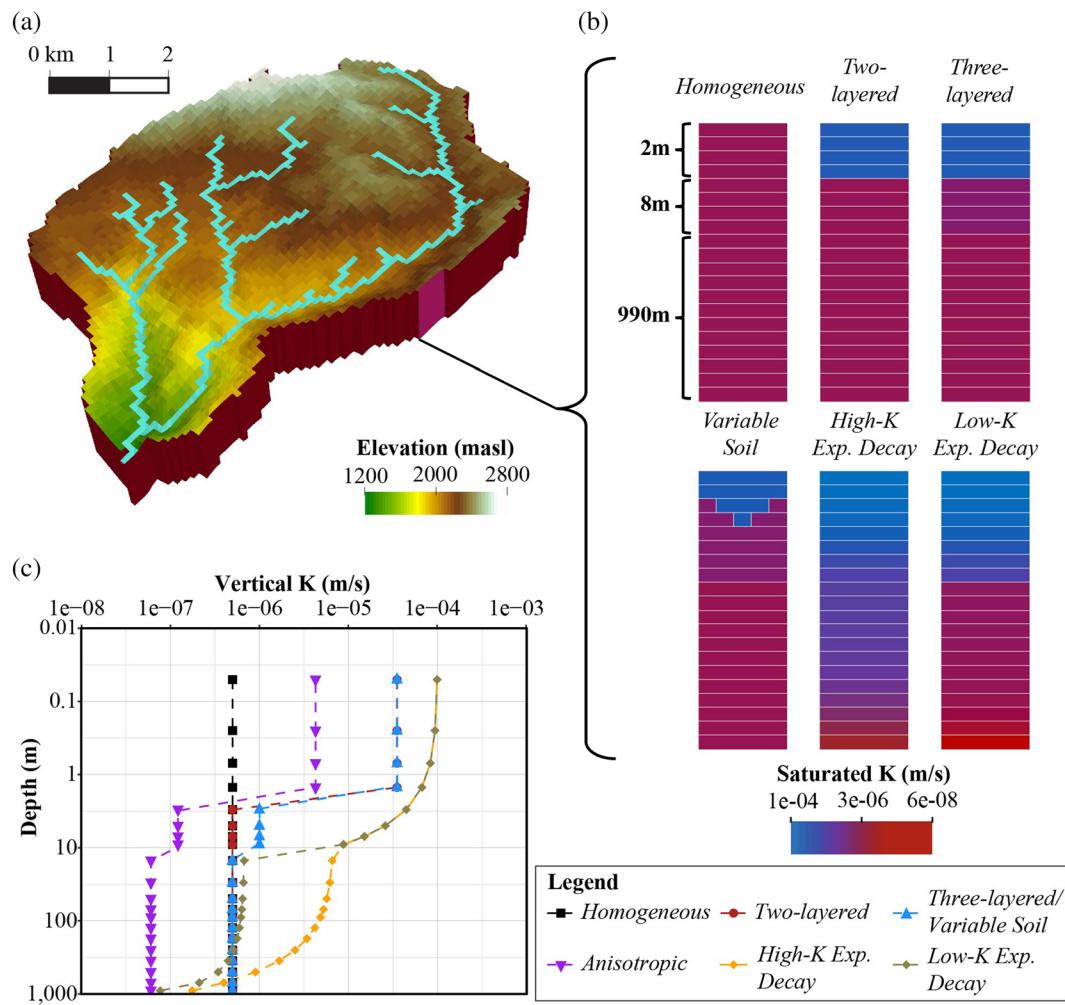
### 2.1.1. Domain Selection and Discretization

We compare scenarios with different subsurface configurations in a semi-idealized domain based on real topography to investigate the sensitivity of flow paths to mountain block heterogeneity. We use topography from a 32-km<sup>2</sup> portion of the headwaters of the Sabino Creek watershed in the Santa Catalina Mountains, north of Tucson, AZ (see Figure 1). The watershed elevation ranges from 2,791 m above sea level (masl) to 1,230 masl. Land cover in this catchment includes subalpine forests, montane fir forests, pine forests, and pine-oak forests (Whittaker & Niering, 1965). The slopes within the catchment range from 0–50%, averaging roughly 22%. We chose this domain for several reasons. First, it includes steepness and topography characteristic of mountain watersheds. Second, we had easily accessible digital elevation model (DEM) and climatic data for the watershed. Finally, this domain is simple enough to see clear flow paths but has complex real-world topography that allows us to see more interesting changes in flow paths when compared to a simple sinusoidal topography that has been used in previous research on idealized groundwater models (Cardenas & Jiang, 2010; Gleeson & Manning, 2008).

To construct the domain, we used 30-m DEM data from the ASTER Global DEM data set (Tachikawa et al., 2011). We upscaled the 30-m DEM to 90 m by choosing the minimum elevations, resulting in 3,948 active grid cells for the domain. Choosing the minimum elevations to aggregate and upscale the DEM data is the best method for capturing the drainage network of the domain. We then processed the upscaled DEM data using PriorityFlow (Condon & Maxwell, 2019) to ensure a hydrologically connected drainage network. The blue shading in Figure 1a shows the expected river drainage network based on calculated drainage areas from the topographic processing. The river network shown here is all cells with a drainage area greater than 60 grid cells (0.49 km<sup>2</sup>). The upscaled DEM data result in a coarser drainage network than exist in the field (i.e., containing less first-order ephemeral streams in its subcatchments). The processed drainage network matches well with the observed drainage network in the basin, although it should be noted that this is a semiarid basin and all of the stream cells shown above do not flow perennially.

Our domain comprises 20 vertical layers of variable thickness, ranging from 0.1 m thick at the top layer to 200 m thick at the bottom layer, resulting in a total thickness of 1 km. We use a terrain-following grid (Maxwell, 2013) approach, so the total thickness is constant across the domain. This layer discretization allows for sufficient freedom to adjust the domain parameters for multiple scenarios. We chose the total depth based on a review of prior literature, which indicates that 1 km is sufficient to capture the full local and intermediate flow paths (Ball et al., 2014; Frisbee et al., 2017). We imposed a no-flow boundary condition on the sides and the bottom of the domain; outflow can only occur via surface water flow and evapotranspiration (ET). Imposing the no-flow boundary condition assumes that the groundwater divide follows the topographic divides.

The assumption of the location of the groundwater divide can affect the distribution of flow paths and water table configurations, especially regarding regional flow paths (Harding, 2012). Welch and Allen (2012) showed that subsurface flux between adjacent mountain subcatchments can change baseflow (which



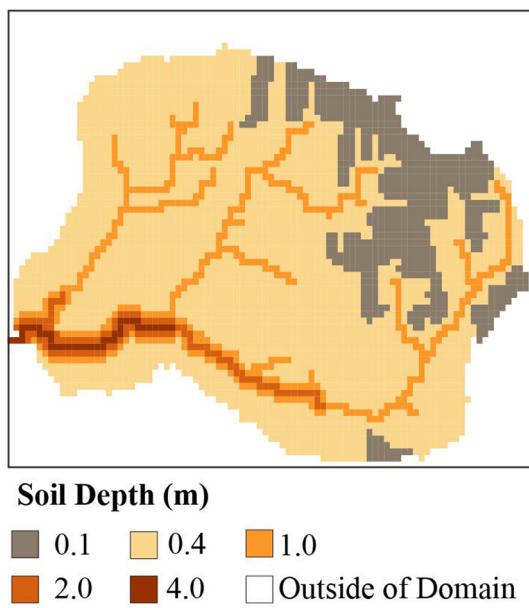
**Figure 1.** (a) The ParFlow-CLM domain, where blue indicates cells in the river network and the remaining color scale indicates elevation. (b) Schematics of the vertical hydraulic conductivity ( $K$ ) values (m/s) of six of the model scenarios (excluding the Anisotropic scenario). Note that the layer thicknesses are not drawn to scale. (c) Plot of the vertical  $K$  values versus depth on a log-log scale. Each point represents the depth at the center of each of the 20 layers.

includes local and intermediate flow paths) up to 10% if there are significant topographic differences between the subcatchments. Our chosen domain is a headwater watershed, and we are assuming here that flow between adjacent basins is insignificant. However, there is potential for what Ameli et al. (2018) describe as “headwater groundwater subsidies” from our watershed to the parent watershed, which follow regional flow paths. We do not have data to support the quantification of these subsidies, and we focus our study on intermediate and local flow paths. Constraining our lateral boundaries to the surface watershed boundary removes this extra degree of freedom and allows a clearer understanding of the response of local and intermediate flow paths to changes in conceptualization of the subsurface layering.

## 2.2. Scenario Design

We developed six scenarios based on the most commonly employed conceptual models in the literature. Figure 1 shows the  $K$  values in each layer for the six scenarios, which are described as follows:

1. Homogeneous (*Homogeneous* scenario): This baseline domain has identical properties throughout. We chose a  $K$  value that is representative of a crystalline bedrock system and experimentally determined a value that produces a plausible water table in our domain. More details follow the scenario descriptions.



**Figure 2.** Plan view of the domain showing the soil depth for the *Variable Soil* scenario. The soil layer has identical properties to the soil layers in the *Two-layered* and *Three-layered* scenarios.

2. Soil and bedrock (*Two-layered* scenario): This domain builds upon the *Homogeneous* scenario by replacing the top 2 m of the domain with a soil layer having a greater  $K$  and porosity. The harmonic mean  $K$  of this scenario is identical to the *Homogeneous* scenario. While the harmonic mean  $K$  will not result in identical solutions to the homogeneous scenario for a variably saturated domain (Zhu, 2008), this is still an appropriate approximation for modeling the broad controls on the hydrology of the mountain block.
3. Soil, saprolite, and bedrock (*Three-layered* scenario): This domain builds upon the *Two-layered* scenario by adding a saprolite layer to the 8 m below the soil layer. The harmonic mean  $K$  of the *Three-layered* scenario is identical to the *Homogeneous* scenario and the *Two-layered* scenario.
4. Variable soil depth (*Variable Soil* scenario): Following some of the common mountain block erosion processes (Pelletier & Rasmussen, 2009), we develop the *Variable Soil* scenario to simulate the effects of spatially varying soil depth based on the distance to a stream cell. The soil depth ranges from 4 to 0.1 m, decreasing with distance from the nearest stream cell, as shown in Figure 2. The maximum soil depth at a river cell also decreases with tributary watershed area, with a minimum thickness of 1 m for any river cell. The mean  $K$  in this scenario is slightly lower than the prior scenarios since the average soil depth was shallower than the *Three-layered* scenario, and we used the same  $K$  values for each of the geologic layers as the *Three-layered* scenario.
5. *Three-layered* scenario with anisotropy (*Anisotropic* scenario): Hydraulic conductivity in mountain blocks can be strongly anisotropic (Mayo et al., 2003), where bedding, weathering processes, and folds produce horizontal  $K$ s that exceed vertical  $K$ s and faults and fractures produce dominant  $K$ s primarily oriented parallel to the dip direction and secondarily to the strike direction (H. Voeckler & Allen, 2012). However, many modeling studies of the mountain block assume isotropic  $K$  (e.g., Gilbert & Maxwell, 2017; Yao et al., 2018). To investigate the effects of anisotropy on our model, we develop the *Anisotropic* scenario by decreasing the magnitude of the vertical  $K_z$  in each layer of the *Three-layered* scenario and increased the magnitude of  $K_x$  and  $K_y$ . We designed this so that the ratio of  $K_x$  and  $K_y$  to  $K_z$  is 10, and the magnitude of the tensor vector is equal to the other scenarios ( $3^{0.5}$ ). This results in multiplying the  $K_x$ ,  $K_y$ , and  $K_z$  from the *Three-layered scenario* by 1.22, 1.22, and 0.122, respectively. The approach maintains a consistent overall  $K$  between scenarios, allowing the model results to be directly comparable. This anisotropy assumes that weathering in fractured crystalline or sedimentary mountain blocks encourages lateral and inhibits vertical flow to a modest degree (Yager et al., 2009).
6. Exponential decay in  $K$  with depth (*Exponential Decay* scenario): Assuming exponential decay in  $K$  with depth is a common conceptual model for flow models in regional groundwater domains and the mountain block (Ameli et al., 2016; Cardenas & Jiang, 2010; Jiang et al., 2009; C. Wang et al., 2018). This exponential decay model was originally introduced by Louis (1972) as an empirical formula fitting in situ permeability tests in fissured rock. We modeled two versions of this scenario for our study: one matching the harmonic mean  $K$  with the *Homogeneous* scenario (*High-K Exponential Decay* scenario) and one reducing the  $K$  of the lower 990 m to achieve a water table that is more comparable with the other five scenarios (*Low-K Exponential Decay* scenario). We discuss the reasoning for these below.

The  $K$  value for the *Homogenous* scenario was determined experimentally by testing a range of realistic average saturated  $K$  values to obtain a steady-state water table and rainfall-runoff ratio that is characteristic of the Sabino Canyon watershed described in the previous section. As noted above, our objective is not to calibrate the model to a specific stream gauge; however, generating a physically plausible water table is necessary to better understand the flow paths within the mountain block. We based our range of simulated  $K$  values on the values used in the studies summarized by Welch and Allen (2014) and the range of  $K$  values for the geology present in crystalline bedrock systems in Freeze and Cherry (1979). The upper and lower bounds of  $K$  in these systems ranges from about  $1 \times 10^{-3}$  m/s for soil layers to about  $1 \times 10^{-8}$  m/s for low- $K$  unweathered

bedrock. Our  $K$  value for the *Homogeneous* scenario is  $5 \times 10^{-7}$  m/s. This is within an order of magnitude of several studies of crystalline bedrock systems similar to the geology of our domain.

The *High-K Exponential Decay* scenario resulted in a water table that was far lower (greater than 100 m deeper on average) than any of the other scenarios (see Figure 9a in section 3.5). This change in water table is a major driver of the changes in discharge patterns and flow paths, which make it more difficult to compare these results adjacent to each other. We therefore included the results of the *High-K Exponential Decay* scenario in a separate section and developed a similar scenario with an exponentially decaying  $K$  that results in a water table that is more comparable to the other five scenarios, constituting the six primary scenarios. For the *Low-K Exponential Decay* scenario, we adjusted the  $K$  in the bottom 990 m to where the  $K$  at the bottom of the domain would equal  $1 \times 10^{-8}$  m/s as averaged across the vertical layers of the domain. The harmonic mean  $K$  for the *Low-K Exponential Decay* scenario is  $2.1 \times 10^{-7}$  m/s.

As is necessary with all models, we made a series of assumptions to constrain our study. We first assumed that the porosity of each of the layers scaled with  $K$ . We kept all of the components of the CLM model identical between the six scenarios. We assumed that the vegetation cover across the entire domain was made up of a mixed forest, which is consistent with the vegetation across most of the domain (see section 2.1.1). Although fractures and other geologic features in the mountain block can alter subsurface flow paths (Welch & Allen, 2014), we assume that each layer can be represented by equivalent porous media. Our goal is to understand the generalized effect of common conceptual models of structural heterogeneity on simulated hydrologic behavior, and the scale of our model domain exceeds the scale of most fractures by several orders of magnitude, so this assumption is justified.

### 2.3. Initialization and Simulation Process

Our simulation process has three main steps: (1) initialization, (2) transient simulation, and (3) particle tracking. In the first step, we initialize each scenario to reach a water table that is in dynamic equilibrium with the input climatic forcings. To accomplish this, we first apply a uniform, constant recharge rate, using ParFlow separately from CLM, to reach a steady-state water table. We determined that steady-state initialization was complete when the change in storage was less than 1% of the recharge. Following this step, we used the coupled ParFlow-CLM and ran multiple 5-year cycles of hourly climate forcing data acquired from the North American Land Data Assimilation System Project Phase 2 (NLDAS-2) (Xia et al., 2012). The NLDAS-2 data have a spatial resolution of one eighth of a degree, and we chose data from the grid cell that covered the center of our domain. We chose the period of water year (WY) 2009 through 2013 since this period was reflective of the long-term precipitation patterns of the region. We determined each scenario to be initialized when the 52-week running average volumes of subsurface storage and surface water discharge at the end of each 5-year cycle varied by less than 1%. Initialization times varied between 20 and 30 model years for each scenario.

After we initialized each of our model scenarios, we simulated our domain using 1 year of transient forcings to generate hourly outputs for EcoSLIM. We chose WY 2009 for this since its precipitation total was the closest to the mean annual precipitation of the 5-year transient initialization period. We then employed EcoSLIM, initializing the domain with 20 particles in each surface cell and running the simulation forward in time. We also tested adding more particles per cell and determined that 20 particles per cell allows us to capture the range of flow path variability while maintaining reasonable simulation times. For each scenario, we ran EcoSLIM using cycles of the 1 year, hourly ParFlow-CLM outputs until at least 92% of active particles exited the domain.

### 2.4. Output Metrics

We focus our analysis on the following model outputs: (1) water table depth, (2) streamflow discharge, (3) particle distributions and depths, and (4) particle age-length relationships. We evaluate the water tables and streamflow discharge to determine how changes in watershed dynamics affect watershed storage and discharge, respectively. We analyze the particle distributions, depths, and age-length relationships to understand how the choice of conceptual model drives the groundwater system dynamics. These dynamics can be important for characterizing subsurface contaminant transport or hydrologic responses on different time scales (Frisbee et al., 2013).

Water tables vary seasonally with our forcings, and the average water table depths presented here are the point where the subsurface storage is closest to the average storage at dynamic equilibrium (i.e., the water table depth most representative of equilibrium). Daily streamflow discharge patterns are calculated by summing hourly discharge. For flow path distributions, probability density functions (PDFs) are calculated by binning the particle saturated ages into 3-year bins. Any particles that are initialized at the surface in EcoSLIM exit immediately. Because the quantity of these particles depends on the extent of ponded area in each scenario, particles that exit within the first year are removed from our analysis. Flow path depths are calculated as the greatest depth below the domain surface that a particle reaches from each origin cell. We calculate the relationship between saturated particle age and length by averaging these values for the exited particles within each origin cell.

Finally, we developed a sensitivity index to quantify the relative sensitivities of each scenario to the primary hydrologic variables. We calculate the sensitivity index as the sum of the absolute value of the percentage change in the mean and the standard deviation of the value of interest (e.g., water table depth, saturated particle age, and discharge) compared to the *Homogeneous* scenario.

### 3. Results

Using the integrated modeling approach and output metrics detailed in section 2, we explore relationships between subsurface conceptualization and simulated depth to the water table (section 3.1), streamflow discharge (section 3.2), flow path ages and depths (section 3.3), and flow path velocities (section 3.4). As noted in section 2.2, we show the results for the *Low-K Exponential Decay* scenario in sections 3.1–3.4 and present the results for the *High-K Exponential Decay* scenario in section 3.5.

#### 3.1. Water Tables

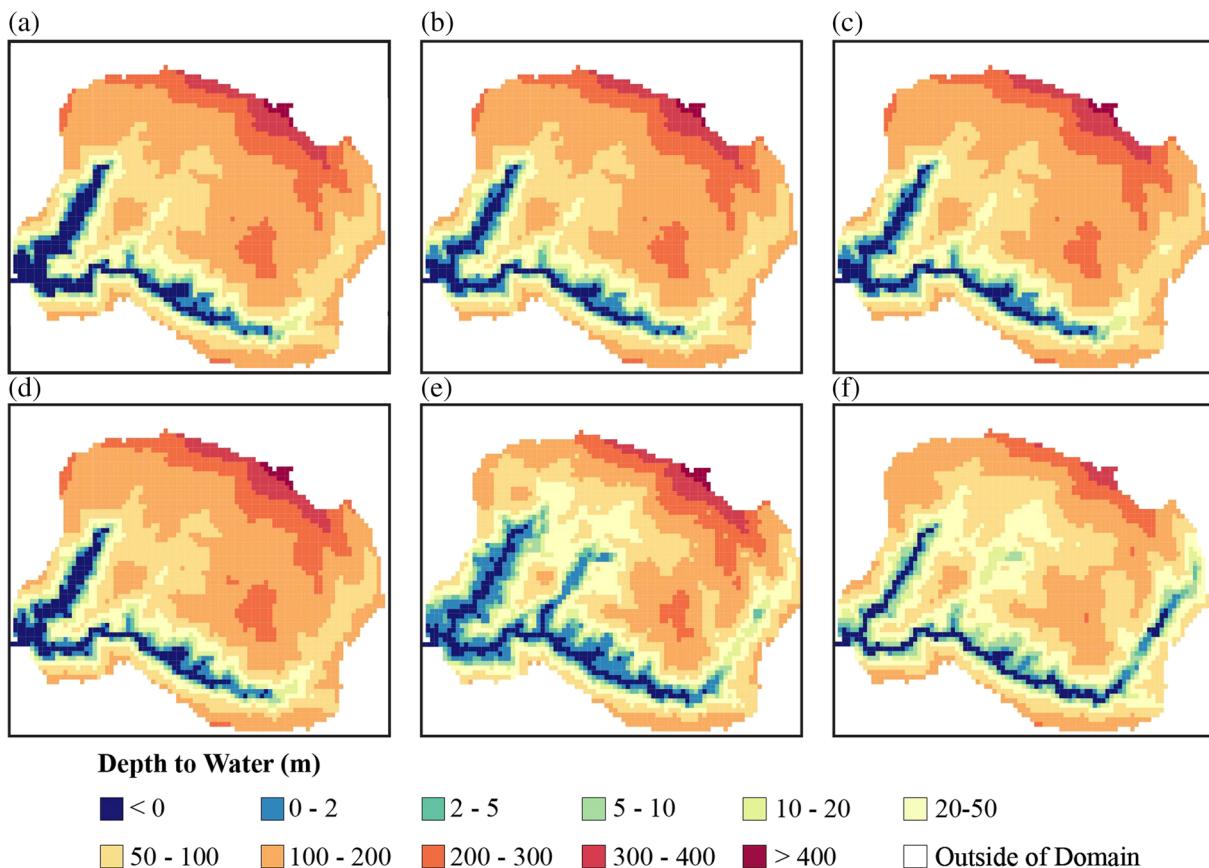
Figures 3a–3f show the water tables for the six primary scenarios. As expected, in all scenarios the water table is the deepest at the highest elevation points in the domain, on the upper right corner of the map, and we see groundwater convergence along the river network and at the outflow in the center of the left side of the map.

The *Homogeneous* scenario results in a water table that intersects the land surface (depth to water less than 0) that follows the rough shape of the expected river network shown in Figure 1 (refer to section 2.1.1 for more details on the river network). Notably, the ponded area (i.e., locations with overland flow) is wider than the processed river network. The ponded areas indicate that local flow paths are not prominent (i.e., discharge is not occurring in domain subcatchments), suggesting that the water table is recharge controlled.

The water tables for the *Two-layered* and the *Three-layered* scenarios are almost identical (Figures 3b and 3c), with a narrower ponded area than the *Homogeneous* scenario. This indicates that the increased  $K$  in the layers of the shallow subsurface results in a slightly lower water table. The similarities in the water tables of the *Two-layered* and the *Three-layered* scenarios suggest that the water table is most sensitive to the addition of the soil layer in the *Two-Layered* scenario and is less sensitive to the addition of the saprolite layer in the *Three-layered* scenario. This is likely due to the relatively small difference in  $K$  below 10 m deep in these scenarios (see Figure 1c). The ponded area in the *Variable Soil* scenario (Figure 3d) is between the *Homogeneous* scenario and the *Two-layered* and *Three-layered* scenarios. This is likely due to the fact that the soil layer is thinner overall in the *Variable Soil* scenario (Figure 2) compared to the *Two-layered* and *Three-layered* scenarios, where the soil layer has a constant thickness of 2 m. In addition, as we kept the total thickness of the soil and saprolite to 10 m in the *Three-layered* and *Variable Soil* scenarios, the *Variable Soil* scenario has a relatively thicker saprolite layer. This change also plays a role in increasing the ponded area in the *Variable Soil* scenario compared to the *Three-layered* scenario.

The *Anisotropic* scenario (Figure 3e) has a shallower water table than the *Three-layered* scenario. The ponded area is also narrower and extends across more of the drainage network than the *Three-layered* scenario (Figure 3c). The area with a water table less than 2 m below the surface is larger in the *Anisotropic* scenario than the other scenarios, which is a reflection of the shallower water table and, along with the narrower ponded area, the increased influence of topography on the water table configuration.

The *Low-K Exponential Decay* scenario (Figure 3f) has a shallower water table on average than the *Homogeneous* scenario due to the lower overall  $K$ . This has a narrower, more extensive ponded area that



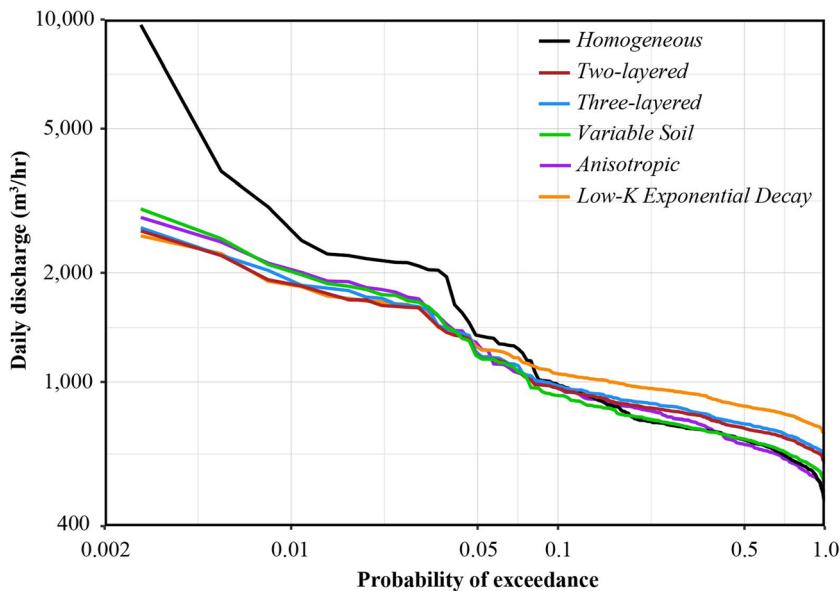
**Figure 3.** Water tables for the six primary scenarios: (a) *Homogeneous*, (b) *Two-layered*, (c) *Three-layered*, (d) *Variable Soil*, (e) *Anisotropic*, and (f) *Low-K Exponential Decay*. These water tables are calculated from the time most representative of the average water table at dynamic equilibrium. The surface topography slopes from the upper right section of the catchment to the center left part of the catchment. Cells that are dark navy indicate water that is ponded on the surface.

more closely resembles the processed river network compared to the other scenarios. The expanded ponded area in the *Low-K Exponential Decay* scenario shows an expansion of local flow paths compared to the *Homogeneous* scenario. The lower overall  $K$  increases the ratio of recharge to  $K$ , which results in a pattern of local flow paths characteristic of a more topography-controlled water table. Compared to the other scenarios, the water table is also shallower at the highest points of the watershed, increasing the hydraulic gradient. This increased gradient highlights the effect of the topography and stream incision on the water table across the domain.

### 3.2. Streamflow Discharge

These systems are in dynamic equilibrium, and the precipitation inputs for each scenario are identical. The total outflows for each scenario are also roughly identical, although each scenario exhibits different partitioning and timing of discharge and ET. Figure 4 shows the flow-duration curves for the six primary scenarios.

The discharge in the *Homogeneous* scenario ranges from 470 to 9,700  $\text{m}^3/\text{hr}$ . All other scenarios have a smaller range in discharge than the *Homogeneous* scenario, with the next highest maximum flow occurring in the *Variable Soil* scenario (3,000  $\text{m}^3/\text{hr}$ ). The decreased discharge rates for high-flow events (probability of exceedance less than 8%) in the heterogeneous scenarios suggest that the higher  $K$  in the shallow subsurface controls the storage-discharge relationship of the watershed. The higher  $K$  layers in the shallow subsurface of the five heterogeneous scenarios increase infiltration, resulting in decreased peaks of overland flow and increased discharge from storage. As precipitation falls on the watershed of the *Homogeneous* scenario, the influx of water exceeds the infiltration capacity of the top layers, resulting in more rapid infiltration-excess type runoff.



**Figure 4.** Flow-duration curves for the six primary scenarios, shown on a log-log scale. These curves are for 1 year of daily discharges aggregated from hourly data.

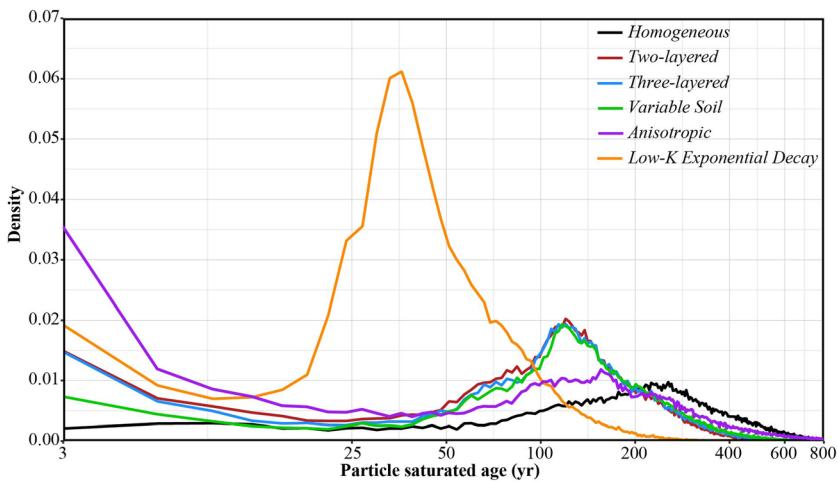
As with groundwater depths, the *Two-layered* and *Three-layered* scenarios have very similar discharge patterns (Figure 4, red and blue, respectively); the median discharge is  $750 \text{ m}^3/\text{hr}$  for the *Two-layered* scenario and  $770 \text{ m}^3/\text{hr}$  for the *Three-layered* scenario, and the maximum flow is  $2,020$  and  $2,050 \text{ m}^3/\text{hr}$  for the *Two-layered* scenario and the *Three-layered* scenario, respectively. The *Variable Soil* scenario, however, exhibits a lower median discharge ( $700 \text{ m}^3/\text{hr}$ ) and a higher range in discharge ( $2,480 \text{ m}^3/\text{hr}$ , defined by the maximum minus the minimum daily flow rate) than the *Two-layered* scenario and the *Three-layered* scenario due to the thinner soils on the hillslopes. The thinner soils reduce the capacity of the hillslopes to capture precipitation falling on the surface.

The hydrograph of the *Anisotropic* scenario (Figure 4, purple) is similar in shape and discharge range to the isotropic scenarios. The maximum discharge of  $2,850 \text{ m}^3/\text{hr}$  lies between the *Variable Soil* scenario and the *Three-layered* scenario. However, the median discharge of the *Anisotropic* scenario is  $670 \text{ m}^3/\text{hr}$ , which is less than the *Two-layered* and *Three-layered* scenarios. The shallower water table in the *Anisotropic* scenario increases the amount of water available for ET, thereby reducing the runoff compared to the isotropic scenarios.

The *Low-K Exponential Decay* scenario's discharge pattern is similar in shape and range to the other scenarios with heterogeneity (Figure 4, orange). Since the overall  $K$  is lower for this scenario, the similar results suggest that the storage-discharge relationship is more sensitive to the  $K$  in the shallow subsurface and not sensitive to the overall  $K$  in the domain past a certain depth. The *Low-K Exponential Decay* scenario has a similar maximum discharge to the other heterogeneous scenarios ( $2,530 \text{ m}^3/\text{hr}$ ), but it has the greatest median discharge of the six primary scenarios ( $860 \text{ m}^3/\text{hr}$ ). The greater magnitude of the low-flow events is due to the smaller ponded area (Figure 3f), resulting in less water exiting through ET.

### 3.3. Flow Path Ages and Depths

PDFs of saturated particle ages illustrate the impact of heterogeneity on flow paths. Figure 5 shows PDFs for the saturated age of exiting particles for the six primary scenarios. Particles in the *Homogeneous* scenario have longer mean saturated ages (282 years) than any of the other scenarios. The mean saturated particle ages for the *Two-* and *Three-layered* scenarios (Figure 5, red and blue, respectively) are 157 and 166 years, respectively. The *Two-* and *Three-layered* scenarios have higher peak density ages than the *Homogeneous* scenario, indicating a concentration of flow paths due to layering. This is consistent with prior modeling research which established that the flow increases and flow paths are concentrated in high- $K$  zones



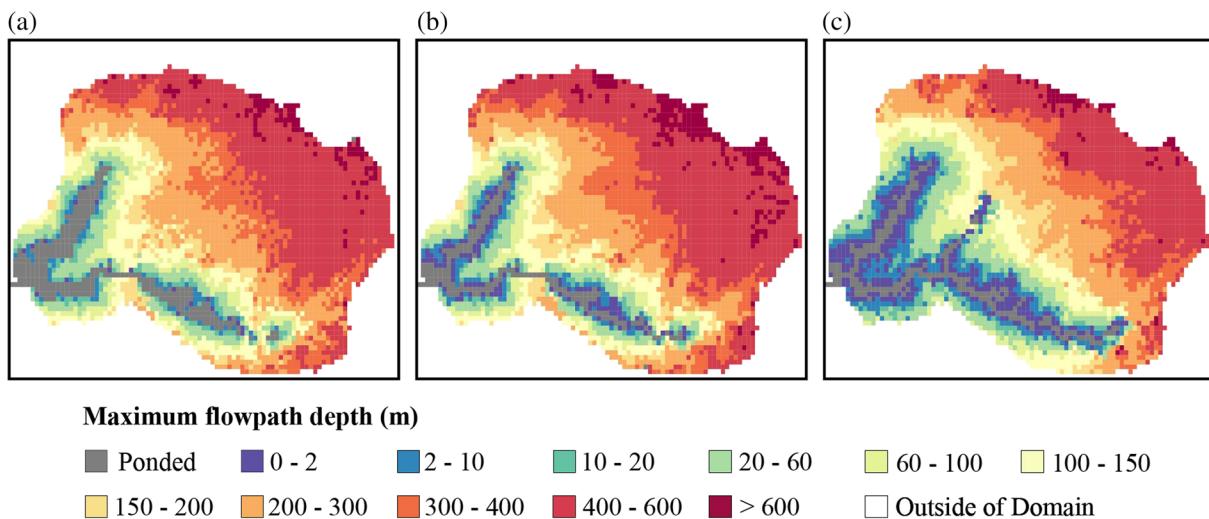
**Figure 5.** Probability density functions (PDFs) of the saturated particle ages for the six primary scenarios, calculated in 3-year bins. These PDFs exclude particles with saturated particle ages of less than 1 year.

(Freeze & Witherspoon, 1967). The *Variable Soil* scenario (Figure 5, green) has flow paths that are less concentrated and older on average than the *Three-layered* scenario (the mean/standard deviation of saturated flow path ages are 166/97 and 181/111 years for the *Three-layered* and *Variable Soil* scenarios, respectively). The differences between the *Two-layered* scenario, *Three-layered* scenario, and the *Variable Soil* scenario are relatively small compared to the difference between these scenarios and the *Homogeneous* scenario. These scenarios all have similar  $K$  values below a depth of 10 m and similar water tables existing primarily more than 10 m below the surface, so we do not expect the saturated flow paths to vary much between them.

The PDF of saturated particle ages for the *Anisotropic* scenario (Figure 5, purple) lies between the *Homogeneous* scenario and the *Two-layered* and *Three-layered* scenarios. The average particle age of the *Anisotropic* scenario is 217 years, which exceeds all scenarios except the *Homogeneous* scenario. The *Anisotropic* scenario spreads out the age distribution, increasing the standard deviation of saturated particle ages to 164 years, the greatest of all six scenarios. Anisotropy has the dual effect of concentrating the flow paths that are more immediately connected with the surface water system while resisting vertical flow from the deeper layers.

The higher water tables in the *Low-K Exponential Decay* scenario (Figure 5, orange) result in younger flow paths on average than the other five scenarios. The mean saturated particle age for the *Low-K Exponential Decay* scenario is 62 years. The higher water table results in a greater portion of flow paths in the layers of the shallow subsurface, which offsets the effect of the lower overall  $K$  and porosity in these scenarios. The *Low-K Exponential Decay* scenario has a high concentration of saturated particle ages indicated by a standard deviation of 47 years, the smallest of the six primary scenarios.

To further understand the changes in flow paths due to anisotropy, we quantified flow path depths for the *Homogeneous* scenario, the *Three-layered* scenario, and the *Anisotropic* scenario (Figures 6a–6c, respectively). Gray cells indicate cells with ponded water where particles exit the domain immediately; these are excluded from our analysis. As expected, flow paths are shallow near the ponded areas and become deeper with increasing distance from a ponded cell. The *Three-layered* scenario has the greatest area with flow paths of greater than 600 m below the surface. The expanded ponded area in the *Anisotropic* scenario results in shallower average flow paths than the *Three-layered* scenario. In addition, the extent of shallow flow paths in the *Anisotropic* scenario is proportionally greater than the *Three-layered* scenario. The *Anisotropic* scenario has 151% more cells with a maximum flow path depth of less than 10 m compared to the *Three-layered* scenario, while only having 16% more cells with ponded water. The *Anisotropic* scenario also has 37% fewer cells with a maximum flow path depth of greater than 400 m compared to the *Three-layered* scenario.



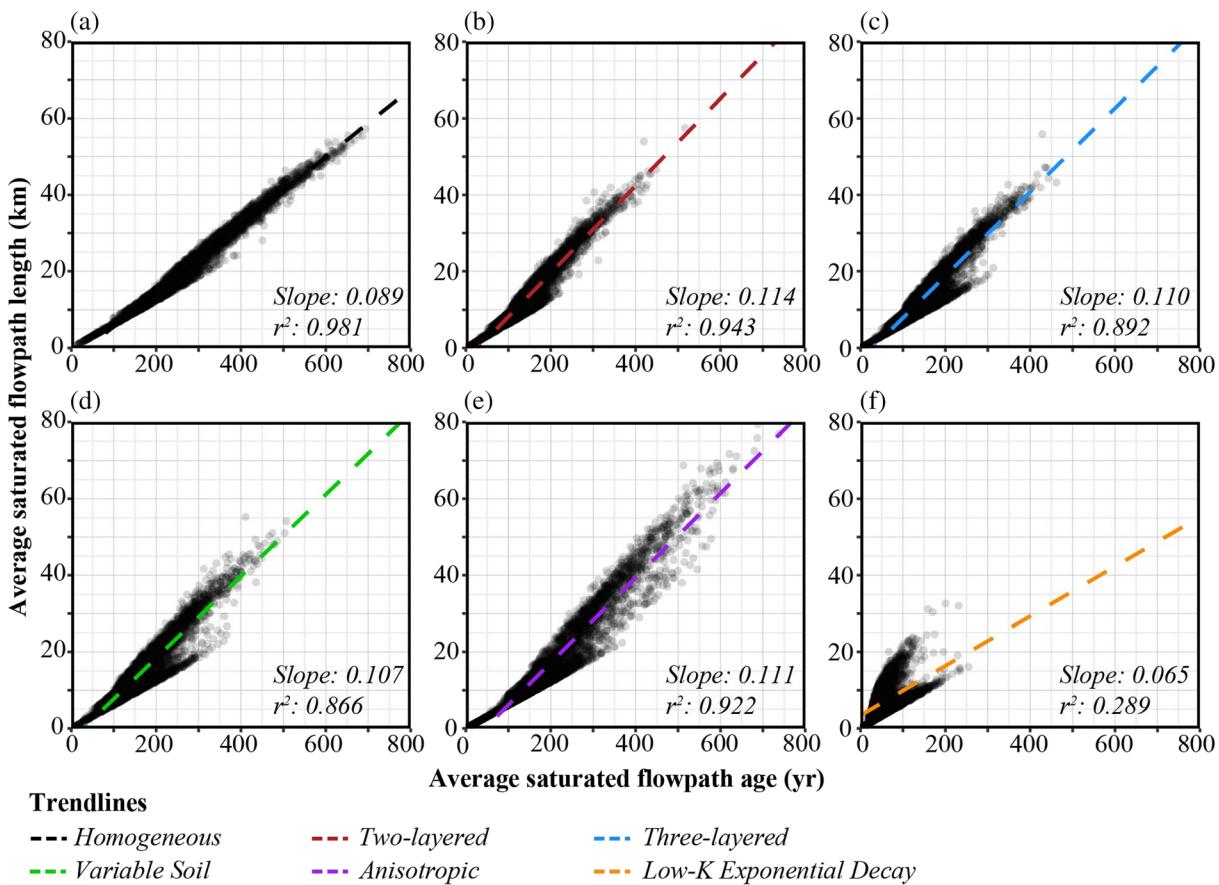
**Figure 6.** Maps of maximum flow path depth below the domain surface for particles at each origin cell for (a) *Homogeneous*, (b) *Three-layered*, and (c) *Anisotropic* scenarios. Particles originating in cells with ponded water (gray cells) exit the domain immediately.

### 3.4. Flow Path Velocities

Comparing the relationship between age and transit length of particles through the saturated zone can help reveal the three-dimensional interactions of topography and heterogeneity on flow paths. The choice of conceptual model drives both the configuration of the water table and the distribution of the particle ages and lengths. As we show in section 3.1 and Figure 3, water table changes can control the extent and location of groundwater discharge points and govern the relative importance of local and intermediate flow paths. For example, in general, we expect that domains with higher  $K$  values to have deeper water tables, resulting in less local flow paths and longer flow paths overall (e.g., comparing the higher- $K$  *Homogeneous* scenario in Figure 3a to the *Low-K Exponential Decay* scenario in Figure 3f). However, longer flow paths do not linearly translate to longer residence times, as a higher  $K$  will also lead to faster velocities (assuming porosity is unchanged). This interplay between flow path length and velocity can be difficult to determine without a model. By comparing the relationships between particle saturated ages and lengths between scenarios, we can better understand the interrelated effects of the water table and conceptual model on flow paths. Figure 7 shows the average age-length relationships of particles for the six primary scenarios. Each plot includes a simple linear trendline to indicate the average relationship between age and path length, where a shallow slope indicates lower velocities and a steeper slope indicates faster velocities. Note that all plots show a bifurcation in the data, where the data with lower length to age ratios (i.e., shallower slopes) characterize the areas where the water table is close to the surface which have negligible vertical flow. The scatter around the trendline of the *Homogeneous* scenario ( $r^2 = 0.981$ ) indicates the effect of topography on velocities, since the hydraulic properties within the domain are homogeneous.

The heterogeneous scenarios have more scatter in the age-length relationship than the *Homogeneous* scenario, as indicated by the lower  $r^2$  values. The *Two-layered* scenario, the *Three-layered* scenario, and the *Variable Soil* scenario (Figures 7b–7d) all have a greater average velocity (linear trendlines with slopes of 0.114, 0.110, and 0.107 km/year, respectively) than the *Homogeneous* scenario (linear trendline slope of 0.089 km/year) due to the concentration of flow paths in the higher- $K$  layers. There are subtle differences between the trendlines in each of these scenarios; the trendline in the *Three-layered* scenario has a shallower slope (0.110 km/year) than the *Two-layered* scenario (0.114 km/year), and the trendline in the *Variable Soil* scenario has a shallower slope (0.107 km/year) than the *Three-layered* scenario.

The *Anisotropic* scenario (Figure 7e) has older particles on average than all other scenarios except for the *Homogeneous* scenario. The trendline for the *Anisotropic* scenario has a slope of 0.111 km/year, which is similar to the slopes of the *Three-layered* and *Variable Soil* scenarios. This suggests that, on average,

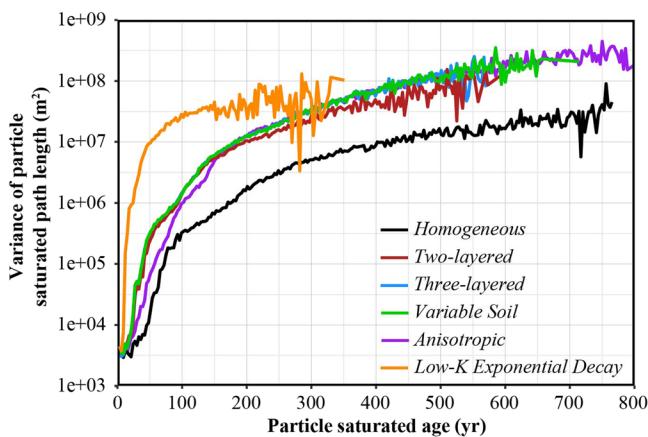


**Figure 7.** Scatter plots of particle saturated age versus particle saturated path length for the six primary scenarios: (a) *Homogeneous*, (b) *Two-layered*, (c) *Three-layered*, (d) *Variable Soil*, (e) *Anisotropic*, and (f) *Low-K Exponential Decay*. Each point represents the average age and flow path length of particles within each starting cell. The dotted lines indicate linear trendlines of saturated age versus saturated flow path length for each scenario, including slopes and  $r$ -squared values.

the higher water table in the *Anisotropic* scenario may counteract the lower vertical  $K$  in the domain compared to the *Three-layered* and *Variable Soil* scenarios.

The scatter plot for the *Low-K Exponential Decay* scenario (Figure 7f) reflects the younger average ages of particles compared to the other scenarios. The *Low-K Exponential Decay* scenario exhibits a clearer bimodal distribution in the particle velocities than the other scenarios, with some cells having velocities closer to the *Homogeneous* scenario and others that have much faster velocities.

To quantify the scatter in the cell-averaged relationship between saturated particle age and particle path length, we calculated the variance of particle path lengths in 1-year age bins. Bins with fewer than 10 particles were excluded from this analysis. The variance plots for the six primary scenarios are shown in Figure 8. For each of the scenarios, the variance in flow path lengths increases as the flow paths increase in age. Similar to the age-length relationships, the heterogeneous scenarios have a greater variance than the *Homogeneous* scenario. The variance in the saturated path length for the *Anisotropic* scenario (Figure 8, purple) is lower than the *Three-layered* scenario at saturated ages of less than 200 years but is similar to the *Three-layered* scenario at older ages. This suggests that anisotropy may not have a significant effect on increasing flow path variance. The *Low-K Exponential Decay* scenario (Figure 8, orange) has the highest variance of saturated flow path lengths overall, resulting from a combination of the greater variability in  $K$ , increased propensity for lateral flow, and the steep vertical gradients of the water table across the domain. Comparing the six primary scenarios suggests that particle path length variance increases with the amount of heterogeneity.



**Figure 8.** Average variance of particle saturated path lengths versus saturated age for the six primary scenarios. Variance of saturated path lengths are calculated in 1-year bins. These plots do not show bins that include fewer than 10 particles.

the *High-K Exponential Decay* scenario (993 m<sup>3</sup>/hr) is greater than all other scenarios. The PDF of particle saturated ages (Figure 9c) shows a peak density of 0.04 greater than the *Low-K Exponential Decay* scenario occurring at an age of less than 25 years. Since the water table is much deeper in the *High-K Exponential Decay* scenario, there are less flow path connections with the land surface, resulting in a greater spatial concentration of flow paths and this change in age distribution. However, more particles exiting the domain at these younger ages results in an increased variability in saturated particle path lengths compared to the *Low-K Exponential Decay* scenario and the *Homogeneous* scenario (Figure 9d).

Finally, the scatter plot of the averaged saturated particle ages and path lengths (Figure 9e) shows a similar bifurcation of the particle velocities. As would be expected with the higher variance in the saturated particle path lengths, the scatter around the trendline is high ( $r^2 = 0.113$ ) compared to the other scenarios. Overall, the layering pattern in the *High-K Exponential Decay* scenario results in a water table that is an outlier compared to the other scenarios, which drives similarly anomalous results in the other hydrologic responses. These results are informative for understanding the implications of the exponentially decaying  $K$  conceptual model of the mountain block subsurface but are less informative to understanding the more nuanced sensitivities of the hydrologic responses to the different conceptual models.

## 4. Discussion

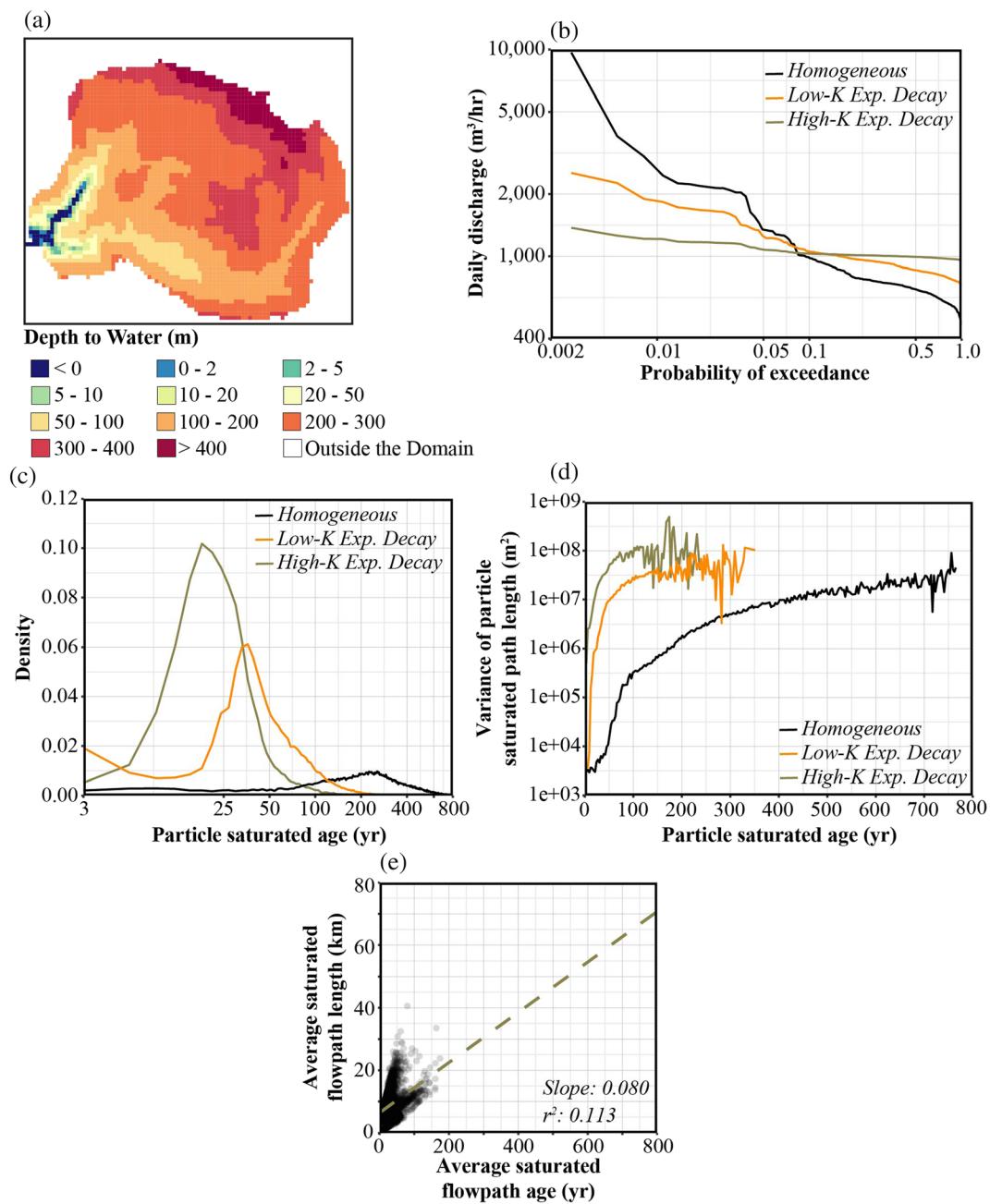
The goal of this research is to characterize the sensitivity of simulated hydrologic behavior in a mountain block watershed to the common choices of subsurface conceptual models. To accomplish this, we use a semi-idealized mountainous domain to simulate scenarios of six common conceptual models of the mountain block. Our combined results reveal important differences in model sensitivity depending on the hydrologic response or characteristic of interest. Figure 10 comprises three subplots comparing the relative sensitivities (described in section 2.4) of the water tables, discharge, and particle age distributions for each scenario compared to the *Homogeneous* scenario. Discharge is sensitive in the six heterogeneous scenarios compared to the *Homogeneous* scenario, but water table is only sensitive to both *Exponential Decay* scenarios and the *Anisotropic* scenario. Particle age distributions, which we use as a proxy for flow paths, are most sensitive to both *Exponential Decay* scenarios, and they are moderately sensitive to the *Two-layered*, *Three-layered*, and *Variable Soil* scenarios. Compared to these scenarios, the greater spread of saturated particle ages in the *Anisotropic* scenario results in a lower flow path sensitivity. All output metrics are most sensitive to the *High-K Exponential Decay* scenario. In the following sections, we discuss these relationships and their drivers in more detail. We discuss the overall impacts of and trade-offs between each conceptual model (section 4.1), which frames a discussion of the implications of our work to modeling mountainous watersheds (section 4.2).

### 3.5. High- $K$ Exponential Decay Scenario Results

Figure 9 shows the results of the *High-K Exponential Decay* scenario, including the water table (Figure 9a), discharge pattern (Figure 9b), PDF of saturated particle ages (Figure 9c), variance of saturated particle lengths (Figure 9d), and a scatter plot of the cell-averaged saturated age versus particle path length. The results of the *Homogeneous* and *Low-K Exponential Decay* scenarios are included in Figures 9b–9d for comparison.

The water table averages over 100 m deeper than the *Homogeneous* scenario and about 150 m deeper than the *Low-K Exponential Decay* scenario. The exponentially decaying  $K$  with depth causes the top portion of the domain to have a much higher  $K$  than the harmonic mean compared to the other scenarios. This results in a water table that appears more recharge controlled than the other scenarios (i.e., less controlled by the topography). The smaller ponded area results in less prominent local flow paths in the domain compared to the other scenarios.

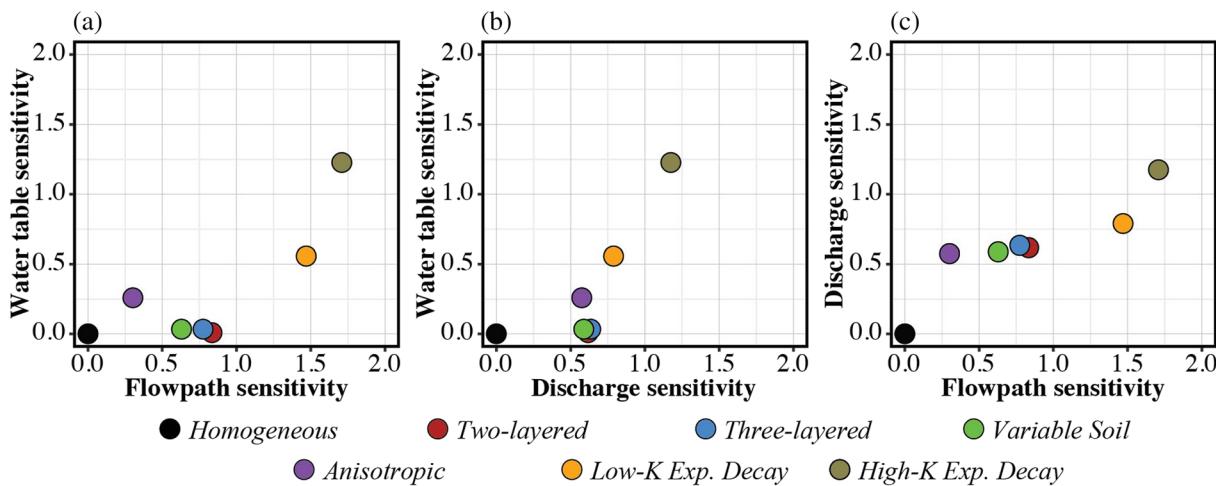
The discharge pattern (Figure 9b) indicates that the *High-K Exponential Decay* scenario is not as sensitive to precipitation as the other scenarios, which follows from the deeper water table. The median discharge for



**Figure 9.** Results of the *High-K Exponential Decay* scenario: (a) water table, (b) flow-duration plot, (c) PDF of saturated particle ages, (d) variance of saturated path lengths, and (e) scatter plot of particle saturated age versus particle saturated path length. Figures 9b–9d include the results of the *Homogeneous* and *Low-K Exponential Decay* scenarios for comparison.

#### 4.1. Conceptual Model Impacts on Simulated Mountain Block Hydrology

First, we consider the impact of different shallow subsurface configurations, specifically the impact of adding higher- $K$  layers to the shallow subsurface. The regional water table configuration is moderately sensitive to the addition of the shallow subsurface layers (Figure 10). The water table is insensitive to spatial variations in soil thickness but is slightly more sensitive to adding a saprolite layer. Adding the saprolite layer, while keeping the harmonic mean  $K$  constant in the domain, resulted in a water table that was slightly shallower than the water tables in the *Homogeneous* and the *Two-layered* scenarios due to the slightly lower  $K$  in the



**Figure 10.** Plots comparing the relative sensitivities of the (a) particle age distributions (used as a proxy for flow paths) and water tables, (b) discharge and water tables, and (c) particle age distributions and discharge for each of the scenarios compared to the *Homogeneous* scenario: All three hydrologic variables are most sensitive to both *Exponential Decay* scenarios. However, the flow paths and discharge patterns are moderately sensitive to the *Two-layered*, *Three-layered*, and *Variable Soil* scenarios.

domain below 10-m depth. Recall that the ponded areas become narrower (Figures 3a–3c) with added high- $K$  layers in the shallow subsurface, which is a surface reflection of the increased capacity of the shallow subsurface to hold water. The higher  $K$  layers in the shallow subsurface strongly control runoff in the catchment by reducing the runoff peaks and increasing baseflow (Figure 4a). Our results show that simply adding a soil layer to an otherwise homogeneous domain can profoundly impact modeled discharge (Figures 10b and 10c). There is a relatively small difference in sensitivity to discharge between the three scenarios with shallow layering and a small increase in sensitivity with the *Low-K Exponential Decay* and *Anisotropic* scenarios (Figure 10). The extent of the regulation of runoff from adding a soil and saprolite layer in the model has important implications for characterizing the fluxes at the land surface and the calibration of watershed models based on hydrographs. For example, many watershed model calibration statistics heavily weight low-frequency, high-flow events (e.g., Nash-Sutcliffe efficiency and mean square error). The differences in the magnitude of the high-flow events for the six primary scenarios (maximum flows ranged from 9,700 m<sup>3</sup>/hr in the *Homogeneous* scenario to 2,530 m<sup>3</sup>/hr in the *Low-K Exponential Decay* scenario) suggest that these calibration metrics are sensitive to changes in the surface properties.

Anisotropy in the mountain block introduces unique changes in the simulated hydrology relative to our other scenarios. Compared to the *Homogeneous* scenario, the *Anisotropic* scenario has a narrower and more expansive ponded area, indicating that topography controls the water table to a greater extent in the *Anisotropic* scenario (Figures 3a and 3e). This follows the behavior predicted by Haitjema and Mitchell-Bruker's (2005) ratio of recharge to  $K$ . Since we decreased the vertical  $K$ , the ratio of recharge to  $K$  increases, and the water table's sensitivity to topography increases. These leads to more local flow paths forming, particularly in a three-dimensional domain. The change in ponded area also occurs in the *Low-K Exponential Decay* scenario (Figure 3f) but to a greater extent because this scenario has a lower overall  $K$  than the *Anisotropic* scenario. The sensitivity of the water table to the *Low-K Exponential Decay* scenario is about double its sensitivity to the *Anisotropic* scenario (Figure 10), suggesting that the addition of exponential decay in  $K$  with depth exerts a stronger control than anisotropy.

The resistance to vertical flow in the *Anisotropic* scenario results in more local flow paths and a shallower hydraulic gradient in the water table, which drives the bifurcation in saturated flow path ages and velocities in our domain (Figure 7e). These results corroborate prior research showing that increasing horizontal anisotropy shortens flow paths in the direction with the higher  $K$  (Gleeson & Manning, 2008). The more extensive ponded area and the shallower water table result in less discharge through streamflow and more discharge through ET from the domain (Figure 4, purple). The isotropic, heterogeneous layering

scenarios all shift the ages, velocities, and discharge patterns in one direction relative to the *Homogeneous* scenario; here we show that anisotropy can affect simulated hydrology by increasing the prevalence of both younger and older flow paths (Figure 5). Our results demonstrate that anisotropy is especially important to incorporate when modeling land surface fluxes, baseflow, and flow path partitioning in the mountain block.

Incorporating spatial variability in soil depth does not change the water table configuration and flow path age distributions in the mountain block very much compared to conceptual models with constant thickness layers. However, there is a notable change in the discharge pattern in the *Variable Soil* scenario compared to the layering scenarios with constant thickness (Figure 4). The reduced infiltration of rainfall in the catchment due to variable soil suggests that this is an important driver of discharge behavior. Although there are multiple conceptual models of soil and weathered bedrock patterns in the mountain block (Rempe & Dietrich, 2014), the scenario that we chose demonstrates the sensitivity of these hydrologic variables to the spatial change in soil depth. Given that the sensitivities of the *Two-layered*, *Three-layered*, and *Variable Soil* scenarios cluster together for all three hydrologic variables (Figure 10), we would expect these sensitivities to be consistent across variations of this conceptual model.

#### 4.2. Implications for Modeling Mountainous Watersheds

In recent years, studies have demonstrated the importance of groundwater flow in mountainous catchments, particularly with respect to baseflow and transit times (Carroll et al., 2019; Condon et al., 2020; Frisbee et al., 2011; Hale & McDonnell, 2016; Rademacher et al., 2005). Various conceptual models will change responses in the hydrologic variables to different extents, and we provide a novel comparison between these variables.

Figure 10 shows that the water table depth is most sensitive to both *Exponential Decay* scenarios, suggesting that the regional water tables in the domain are sensitive to the bulk  $K$  of the domain and the distribution of  $K$  within the domain. We also found that the representation of the river network as ponded area is most sensitive to the  $K$  distribution in the upper layers. Therefore, if the regional water table is the primary variable of interest, assuming a homogeneous  $K$  may be valid. However, if more localized surface water and groundwater interactions are of interest, as is the scenario with riparian behavior, it is important to represent the heterogeneity in the shallow subsurface (Katsuyama et al., 2005).

Unlike water tables, discharge patterns are more sensitive to the  $K_s$  in the shallow subsurface, and our results suggest that heterogeneity at depths past 10 to 100 m does not significantly impact results. This is in line with the concept of an active layer depth approach in mountain block studies. While many definitions exist for what constitutes active circulation, studies in fractured crystalline mountain block systems have roughly converged on a  $K$  value of  $10^{-8}$  m/s (typically 100–200 m deep) beyond which groundwater flow is negligible (Markovich et al., 2019). The extent of the active layer varies based on the bedrock type and degree of weathering, faulting, and fracturing but in general will be shallower in crystalline mountain blocks and deeper in sedimentary and volcanic mountain blocks (Frisbee et al., 2017; Manning & Solomon, 2005). More work is needed to characterize this depth in mountain block systems, particularly given its usefulness as a simplifying assumption for simulating mountain block discharge.

Our choice of a 1-km model domain depth also affected these results. Although our results show that discharge was not sensitive to heterogeneity at depths past 10 to 100 m, particle flow paths reached depths of greater than 600 m below the ground surface, even in the *Anisotropic* scenario (Figure 6). The PDF of saturated particle ages (Figure 5) and the scatter plots (Figure 7) show the particles that reach these depths have a significant vertical flow component, generating deeper and longer flow paths with greater average velocities. By quantifying the cumulative statistics of flow paths for each cell across the watershed, we can discern the salient controls on the flow paths in our domain. This emphasizes the benefit of modeling mountain topography in three dimensions. Our domain depth allowed us to capture these longer intermediate flow paths that affect these flow path results, which would have likely been missed if we would have restricted the domain depth to the shallower layers of active circulation.

The choice of exponentially decaying  $K$  with depth significantly impacts the hydrologic response. Every output metric is most sensitive to the *High- $K$  Exponential Decay* scenario (Figure 10), where the harmonic mean  $K$  matches the  $K$  of the *Homogeneous* scenario. C. Wang et al. (2018) modeled multiple scenarios of an

exponentially decaying  $K$  with depth and showed that, even if the water table is specified and does not change between scenarios, hydrologic responses are still somewhat sensitive to the distribution of  $K$ . Our results build upon this to relax the assumption of the water table location, further demonstrating the importance of choosing the appropriate conceptual model.

While total discharge is relatively insensitive to the choice of conceptual model below some depth, if results depend on groundwater velocity or residence time distributions, our results suggest that both shallow high- $K$  layers and overall hydraulic conductivity can be important. Simply adding a shallow layer of soil, as in the *Two-layered* scenario, altered the average velocity and the distribution of saturated particle ages, resulting in a flow path sensitivity of 0.83 (Figure 10). The sensitivity of discharge to the *Low- $K$  Exponential Decay* scenario (0.79) is near the *Two-layered* scenario (0.68), but the flow path sensitivities are greater (1.47 and 0.83 for the *Low- $K$  Exponential Decay* scenario and the *Two-layered* scenario, respectively).

Similarly, the variance in the age-length relationship of particle path lengths increases with the extent of heterogeneity in the domain (Figures 8 and 9d). Notably, adding shallow high- $K$  zones decreased the mean particle age by a factor of 1.8 and adding exponential decay further decreased the mean particle ages by a factor of 2.5 (comparing the *Two-layered* scenario to the *Low- $K$  Exponential Decay* scenario). The decrease in mean particle ages from the *Two-layered* scenario to the *Low- $K$  Exponential Decay* scenario shows the interplay between the lower overall  $K$  and shallower water table in the *Low- $K$  Exponential Decay* scenario. This finding has implications for using hydrologic models to characterize groundwater age distributions in mountainous watersheds (Engdahl & Maxwell, 2015; Maxwell et al., 2016) as well as for using environmental tracer data to calibrate hydrologic models (Sanford, 2011). Calibrating models to tracer data has been shown to be a powerful tool for reducing parameter nonuniqueness (Schilling et al., 2019); however, our findings point to the need to first evaluate the potentially large effect conceptual assumptions of the mountain block may have on groundwater age distributions. Understanding the degree to which heterogeneity decreases mean age and increases age variance can improve our ability to characterize uncertainty with respect to groundwater age distributions in the mountain block.

Our results are also a product of the horizontal resolution of our domain. Flow path distributions and discharge patterns can vary with the domain resolution (C. Wang et al., 2018). A higher-resolution simulation would explicitly represent more small ephemeral streams and capture more local flow paths. The sensitivity of the hydrologic response to the domain resolution increases with increased topographic relief. We chose a horizontal domain resolution of 90 m to balance the computational demand with topographic complexity in our watershed and because we think this is representative of other watershed-scale models. Our water table results (Figure 3) indicate that the water tables are primarily recharge controlled, with less prominent local flow paths. Therefore, the flow path partitioning will be less sensitive to the domain resolution. Furthermore, while the measured hydrologic responses will change with varying resolutions, the relative effects of the choice of conceptual model remain the same (C. Wang et al., 2018).

Finally, the effect of the subsurface conceptualization on the simulated hydrology depends on the scale of interest. We chose our study domain extents to capture local and intermediate flow paths in a subcatchment of a larger mountain block, excluding regional flow paths. Modeling in domains that capture regional flow indicates that regional flow paths are sensitive to water tables and other hydrologic metrics (Gleeson & Manning, 2008), suggesting that our results are applicable to understanding the implications of conceptual models in larger watersheds. While we constrained our boundaries to a mountain subcatchment, these sensitivities are applicable to understanding local and intermediate flow paths in other catchments. Sensitivity studies over larger, regional watersheds with less steep topography (Goderniaux et al., 2013) show that areas with local and intermediate flow paths exhibit smaller variations with changes in recharge when compared to the longer, regional flow paths. Goderniaux et al. (2013) also indicate similar shapes of the local and intermediate flow paths over a wide range of recharge rates, suggesting that the sensitivities of flow paths to the conceptual model may be muted with less steep topography. This demonstrates the applicability of our findings to nonmountainous systems, as the primary parameters affecting regional groundwater systems are  $K$ , relief, and recharge. Settings with lower relief increase the relative influence of the choice of conceptual model ( $K$  varying with depth) since relief is less prominent, assuming constant recharge. Compared with several prior sensitivity studies (Gleeson & Manning, 2008; Goderniaux et al., 2013), we further demonstrate that the salient controls on groundwater systems depend on the domain. Our approach identifies these

controls on a high-relief mountainous groundwater system to further inform the design of groundwater models for areas with similar characteristics.

## 5. Conclusions

This study compares the simulated hydrologic behavior resulting from six common conceptual models of the mountain block subsurface. These include multiple representations of hydraulic conductivity ( $K$ ) varying with depth, changes in soil depth with topography, and anisotropic geology. We use ParFlow-CLM coupled with EcoSLIM to evaluate flow paths, discharge, and water tables to quantify the impact of subsurface conceptualization on simulated hydrologic behavior.

We show that the sensitivity of simulated hydrology to the choice of conceptual model varies noticeably depending on the variable of interest. For example, we show that water table configuration is not very sensitive to shallow layering if the overall  $K$  in the domain is the same. We only see significant changes in the water table with changes in  $K$  distribution at greater depths (as in the *High-K Exponential Decay* scenario, Figure 9a). However, these same differences in conceptual model have a much larger impact on the simulated discharge. We also demonstrate that the homogeneous domain results in flow paths that are deeper and discharge that is more sensitive to precipitation than domains with higher- $K$  layers in the shallow subsurface. The conceptual model with exponentially decaying  $K$  with depth results in a more well-defined area where the water table intersects the land surface. Anisotropy produces flow paths that are shallower and older on average than isotropic scenarios and restricts the discharge from storage. Overall, layering increases the variability of subsurface flow paths in the mountain block.

This research builds on simpler idealized modeling that has formed our basic understanding of trade-offs between water tables, flow paths, recharge, and discharge in both regional and mountain block groundwater systems, generally in idealized settings (Gleeson & Manning, 2008; Haitjema & Mitchell-Bruker, 2005; Tóth, 1963; Welch & Allen, 2012). We expand on this work focusing specifically on mountain block hydrology and more complex real topography, which poses unique challenges for both observations and modeling. Although we only chose six conceptualizations of a mountain block imposed on one topographic domain, we used simple and commonly applied conceptual models to produce results that can be relevant to other mountain block watersheds. Our results showing the cumulative hydrologic responses and changes in data distribution across our watershed highlight the importance of modeling complex topography in three dimensions. Understanding the salient drivers of changes in the simulated mountain block hydrology can inform the development of future models of mountain block watersheds and help guide data collection to constrain the parameters of a mountain block hydrology model.

The analysis we present provides insights into how watershed modeling choices may bias the conclusions drawn from different simulation approaches. We show that discharge, water tables, and flow paths respond differently to the choice of conceptual model, and evaluating a model on one metric risks biasing the model outputs, limiting the model's usefulness to interpreting uncertain field measurements or making predictions. Understanding these trade-offs and potential biases is vital for characterizing the potential future changes to mountain water resources.

## Conflict of Interest

The authors declare no conflicts of interest.

## Data Availability Statement

The ParFlow-CLM and EcoSLIM input scripts and outputs used for this research are available online (<https://doi.org/10.5281/zenodo.4009879>).

## References

- Ajami, H., Troch, P. A., Maddock, T., Meixner, T., & Eastoe, C. (2011). Quantifying mountain block recharge by means of catchment-scale storage-discharge relationships. *Water Resources Research*, 47, W04504. <https://doi.org/10.1029/2010WR009598>
- Ameli, A. A., Gabrielli, C., Morgenstern, U., & McDonnell, J. J. (2018). Groundwater subsidy from headwaters to their parent water watershed: A combined field-modeling approach. *Water Resources Research*, 54, 5110–5125. <https://doi.org/10.1029/2017WR022356>

## Acknowledgments

This research was supported by the National Science Foundation (NSF) Postdoctoral Fellowship from the Division of Earth Sciences (EAR-1806383). The University of Arizona's High-Performance Computing (UAHPC) team provided resources to facilitate the model runs.

- Ameli, A. A., McDonnell, J. J., & Bishop, K. (2016). The exponential decline in saturated hydraulic conductivity with depth: A novel method for exploring its effect on water flow paths and transit time distribution. *Hydrological Processes*, 30, 2438–2450. <https://doi.org/10.1002/hyp.10777>
- Ashby, S. F., & Falgout, R. D. (1996). A parallel multigrid preconditioned conjugate gradient algorithm for groundwater flow simulations. *Nuclear Science and Engineering*, 124(1), 145–159. <https://doi.org/10.13182/NSE96-A24230>
- Ball, L. B., Caine, J. S., & Ge, S. (2014). Controls on groundwater flow in a semiarid folded and faulted intermountain basin. *Water Resources Research*, 50, 6788–6809. <https://doi.org/10.1002/2013WR014451>
- Cardenas, M. B., & Jiang, X.-W. (2010). Groundwater flow, transport, and residence times through topography-driven basins with exponentially decreasing permeability and porosity. *Water Resources Research*, 46, W11538. <https://doi.org/10.1029/2010WR009370>
- Carroll, R. W. H., Deems, J. S., Niswonger, R., Schumer, R., & Williams, K. H. (2019). The importance of interflow to groundwater recharge in a snowmelt-dominated headwater basin. *Geophysical Research Letters*, 46, 5899–5908. <https://doi.org/10.1029/2019GL082447>
- Condon, L. E., Markovich, K. H., Kelleher, C. A., McDonnell, J. J., Ferguson, G., & McIntosh, J. C. (2020). Where is the bottom of a watershed? *Water Resources Research*, 56, e2019WR026010. <https://doi.org/10.1029/2019WR026010>
- Condon, L. E., & Maxwell, R. M. (2019). Modified priority flood and global slope enforcement algorithm for topographic processing in physically based hydrologic modeling applications. *Computers & Geosciences*, 126, 73–83. <http://www.sciencedirect.com/science/article/pii/S0098300418307957>, <https://doi.org/10.1016/j.cageo.2019.01.020>
- Engdahl, N. B., & Maxwell, R. M. (2015). Quantifying changes in age distributions and the hydrologic balance of a high-mountain watershed from climate induced variations in recharge. *Journal of Hydrology*, 522, 152–162. <http://www.sciencedirect.com/science/article/pii/S0022169414010439>, <https://doi.org/10.1016/j.jhydrol.2014.12.032>
- Forster, C., & Smith, L. (1988). Groundwater flow systems in mountainous terrain: 2. Controlling factors. *Water Resources Research*, 24(7), 1011–1023. <https://doi.org/10.1029/WR024i007p01011>
- Freeze, R. A., & Cherry, J. (1979). *Groundwater*. Englewood Cliffs: Prentice-Hall.
- Freeze, R. A., & Witherspoon, P. A. (1967). Theoretical analysis of regional groundwater flow: 2. Effect of water-table configuration and subsurface permeability variation. *Water Resources Research*, 3(2), 623–634. <https://doi.org/10.1029/WR003i002p00623>
- Frisbee, M. D., Phillips, F. M., Campbell, A. R., Liu, F., & Sanchez, S. A. (2011). Streamflow generation in a large, alpine watershed in the southern Rocky Mountains of Colorado: Is streamflow generation simply the aggregation of hillslope runoff responses? *Water Resources Research*, 47, W06512. <https://doi.org/10.1029/2010WR009391>
- Frisbee, M. D., Tolley, D. G., & Wilson, J. L. (2017). Field estimates of groundwater circulation depths in two mountainous watersheds in the western U.S. and the effect of deep circulation on solute concentrations in streamflow. *Water Resources Research*, 53, 2693–2715. <https://doi.org/10.1002/2016WR019553>
- Frisbee, M. D., Wilson, J. L., Gomez-Velez, J. D., Phillips, F. M., & Campbell, A. R. (2013). Are we missing the tail (and the tale) of residence time distributions in watersheds? *Geophysical Research Letters*, 40, 4633–4637. <https://doi.org/10.1002/grl.50895>
- Gilbert, J. M., & Maxwell, R. M. (2017). Examining regional groundwater-surface water dynamics using an integrated hydrologic model of the San Joaquin River basin. *Hydrology and Earth System Sciences*, 21, 923–947. <https://doi.org/10.5194/hess-21-923-2017>
- Gleeson, T., & Manning, A. H. (2008). Regional groundwater flow in mountainous terrain: Three-dimensional simulations of topographic and hydrogeologic controls. *Water Resources Research*, 44, W10403. <https://doi.org/10.1029/2008WR006848>
- Goderniaux, P., Davy, P., Bresciani, E., Dreuzey, J.-R., & Borgne, T. (2013). Partitioning a regional groundwater flow system into shallow local and deep regional flow compartments. *Water Resources Research*, 49, 2274–2286. <https://doi.org/10.1002/wrcr.20186>
- Haitjema, H. M., & Mitchell-Bruker, S. (2005). Are water tables a subdued replica of the topography? *Ground Water*, 43(6), 781–786. <https://www.ncbi.nlm.nih.gov/pubmed/16323999>, <https://doi.org/10.1111/j.1745-6584.2005.00090.x>
- Hale, V. C., & McDonnell, J. J. (2016). Effect of bedrock permeability on stream base flow mean transit time scaling relations: 1. A multiscale catchment intercomparison. *Water Resources Research*, 52, 1358–1374. <https://doi.org/10.1002/2014WR016124>
- Harding, J. (2012). *Building the foundation for a hydrogeology model of the Rio Hondo Watershed, NM* (Master's thesis). Retrieved from New Mexico Bureau of Geology and Mineral Resources ([http://www.ees.nmt.edu/outside/alumni/papers/2012t\\_harding\\_jj.pdf](http://www.ees.nmt.edu/outside/alumni/papers/2012t_harding_jj.pdf)). Socorro, NM: New Mexico Institute of Mining and Technology.
- Immerzeel, W. W., Lutz, A. F., Andrade, M., Bahl, A., Biemans, H., Bolch, T., et al. (2019). Importance and vulnerability of the world's water towers. *Nature*, 577, 364–369. <https://doi.org/10.1038/s41586-019-1822-y>
- Jiang, X.-W., Wan, L., Wang, X.-S., Ge, S., & Liu, J. (2009). Effect of exponential decay in hydraulic conductivity with depth on regional groundwater flow. *Geophysical Research Letters*, 36, L24402. <https://doi.org/10.1029/2009GL041251>
- Jones, J. E., & Woodward, C. S. (2001). Newton-Krylov-multigrid solvers for large-scale, highly heterogeneous, variably saturated flow problems. *Advances in Water Resources*, 24(7), 763–774. <http://www.sciencedirect.com/science/article/pii/S0309170800000750>, [https://doi.org/10.1016/S0309-1708\(00\)00075-0](https://doi.org/10.1016/S0309-1708(00)00075-0)
- Katsuyama, M., Ohte, N., & Kabeya, N. (2005). Effects of bedrock permeability on hillslope and riparian groundwater dynamics in a weathered granite catchment. *Water Resources Research*, 41, W01010. <https://doi.org/10.1029/2004WR003275>
- Kollet, S. J., & Maxwell, R. M. (2006). Integrated surface-groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. *Advances in Water Resources*, 29(7), 945–958. <http://www.sciencedirect.com/science/article/pii/S0309170805002101>, <https://doi.org/10.1016/j.advwatres.2005.08.006>
- Kollet, S. J., & Maxwell, R. M. (2008). Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model. *Water Resources Research*, 44, W02402. <https://doi.org/10.1029/2007WR006004>
- Louis, C. (1972). Rock hydraulics. In L. Müller (Ed.), *Rock mechanics* (pp. 299–387). Vienna: Springer Vienna. <https://doi.org/10.1007/978-3-7091-4109-016>
- Manning, A. H., & Solomon, D. K. (2005). An integrated environmental tracer approach to characterizing groundwater circulation in a mountain block. *Water Resources Research*, 41, W12412. <https://doi.org/10.1029/2005WR004178>
- Markovich, K. H., Manning, A. H., Condon, L. E., & McIntosh, J. C. (2019). Mountain-block recharge: A review of current understanding. *Water Resources Research*, 55, 8278–8304. <https://doi.org/10.1029/2019WR025676>
- Markovich, K. H., Maxwell, R. M., & Fogg, G. E. (2016). Hydrogeological response to climate change in alpine hillslopes. *Hydrological Processes*, 30(18), 3126–3138. <https://doi.org/10.1002/hyp.10851>
- Maxwell, R. M. (2013). A terrain-following grid transform and preconditioner for parallel, large-scale, integrated hydrologic modeling. *Advances in Water Resources*, 53, 109–117. <http://www.sciencedirect.com/science/article/pii/S0309170812002564>, <https://doi.org/10.1016/j.advwatres.2012.10.001>

- Maxwell, R. M., Condon, L. E., Danesh-Yazdi, M., & Bearup, L. A. (2019). Exploring source water mixing and transient residence time distributions of outflow and evapotranspiration with an integrated hydrologic model and Lagrangian particle tracking approach. *Ecohydrology*, 12, e2042. <https://doi.org/10.1002/eco.2042>
- Maxwell, R. M., Condon, L. E., Kollet, S. J., Maher, K., Haggerty, R., & Forrester, M. M. (2016). The imprint of climate and geology on the residence times of groundwater. *Geophysical Research Letters*, 43, 701–708. <https://doi.org/10.1002/2015GL066916>
- Maxwell, R. M., & Miller, N. L. (2005). Development of a coupled land surface and groundwater model. *Journal of Hydrometeorology*, 6(3), 233–247. <https://doi.org/10.1175/JHM422.1>
- Mayo, A. L., Morris, T. H., Peltier, S., Petersen, E. C., Payne, K., Holman, L. S., et al. (2003). Active and inactive groundwater flow systems: Evidence from a stratified, mountainous terrain. *GSA Bulletin*, 115(12), 1456–1472. <https://doi.org/10.1130/B25145.1>
- Pelletier, J. D., & Rasmussen, C. (2009). Geomorphically based predictive mapping of soil thickness in upland watersheds. *Water Resources Research*, 45, W09417. <https://doi.org/10.1029/2008WR007319>
- Rademacher, L. K., Clark, J. F., Clow, D. W., & Hudson, G. B. (2005). Old groundwater influence on stream hydrochemistry and catchment response times in a small Sierra Nevada catchment: Sagehen Creek, California. *Water Resources Research*, 41, W02004. <https://doi.org/10.1029/2003WR002805>
- Rempe, D. M., & Dietrich, W. E. (2014). A bottom-up control on fresh-bedrock topography under landscapes. *Proceedings of the National Academy of Sciences*, 111, 6576, 6581. <http://www.pnas.org/content/111/18/6576.abstract>, <https://doi.org/10.1073/pnas.1404763111>
- Sanford, W. (2011). Calibration of models using groundwater age. *Hydrogeology Journal*, 19(1), 13–16. <https://doi.org/10.1007/s10040-010-0637-6>
- Schilling, O. S., Cook, P. G., & Brunner, P. (2019). Beyond classical observations in hydrogeology: The advantages of including exchange flux, temperature, tracer concentration, residence time, and soil moisture observations in groundwater model calibration. *Reviews of Geophysics*, 57(1), 146–182. <https://doi.org/10.1029/2018RG000619>
- Tachikawa, T., Kaku, M., Iwasaki, A., Gesch, D. B., Oimoen, M. J., Zhang, Z., et al. (2011). ASTER global digital elevation model version 2—Summary of validation results. Retrieved from <http://pubs.er.usgs.gov/publication/70005960>
- Tetzlaff, D., Buttle, J., Carey, S. K., McGuire, K., Laudon, H., & Soulsby, C. (2015). Tracer-based assessment of flow paths, storage and runoff generation in northern catchments: A review. *Hydrological Processes*, 29, 3475–3490. <https://doi.org/10.1002/hyp.10412>
- Tóth, J. (1963). A theoretical analysis of groundwater flow in small drainage basins. *Journal of Geophysical Research*, 68(16), 4795–4812. <https://doi.org/10.1029/JZ068i016p04795>
- Voeckler, H., & Allen, D. M. (2012). Estimating regional-scale fractured bedrock hydraulic conductivity using discrete fracture network (DFN) modeling. *Hydrogeology Journal*, 20(6), 1081–1100. <https://doi.org/10.1007/s10040-012-0858-y>
- Voeckler, H. M., Allen, D. M., & Alila, Y. (2014). Modeling coupled surface water—Groundwater processes in a small mountainous headwater catchment. *Journal of Hydrology*, 517, 1089–1106. <http://www.sciencedirect.com/science/article/pii/S0022169414004752>, <https://doi.org/10.1016/j.jhydrol.2014.06.015>
- Wang, C., Gomez-Velez, J. D., & Wilson, J. L. (2018). The importance of capturing topographic features for modeling groundwater flow and transport in mountainous watersheds. *Water Resources Research*, 54, 10, 313–310,338. <https://doi.org/10.1029/2018WR023863>
- Wang, J.-Z., Jiang, X.-W., Zhang, Z.-Y., Wan, L., Wang, X.-S., & Li, H. (2017). An analytical study on three-dimensional versus two-dimensional water table-induced flow patterns in a Tóthian basin. *Hydrological Processes*, 31, 4006–4018. <https://doi.org/10.1002/hyp.11317>
- Welch, L. A., & Allen, D. M. (2012). Consistency of groundwater flow patterns in mountainous topography: Implications for valley bottom water replenishment and for defining groundwater flow boundaries. *Water Resources Research*, 48, W05526. <https://doi.org/10.1029/2011WR010901>
- Welch, L. A., & Allen, D. M. (2014). Hydraulic conductivity characteristics in mountains and implications for conceptualizing bedrock groundwater flow. *Hydrogeology Journal*, 22, 1003–1026. <https://doi.org/10.1007/s10040-014-1121-5>
- Whittaker, R. H., & Niering, W. A. (1965). Vegetation of the Santa Catalina Mountains, Arizona: A gradient analysis of the south slope. *Ecology*, 46(4), 429–452. [www.jstor.org/stable/1934875](http://www.jstor.org/stable/1934875), <https://doi.org/10.2307/1934875>
- Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., et al. (2012). Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLadas-2): 1. Intercomparison and application of model products. *Journal of Geophysical Research*, 117, D03109. <https://doi.org/10.1029/2011JD016048>
- Yager, R. M., Voss, C. I., & Southworth, S. (2009). Comparison of alternative representations of hydraulic-conductivity anisotropy in folded fractured-sedimentary rock: Modeling groundwater flow in the Shenandoah Valley (USA). *Hydrogeology Journal*, 17(5), 1111–1131. <https://doi.org/10.1007/s10040-008-0431-x>
- Yao, Y., Tian, Y., Andrews, C., Li, X., Zheng, Y., & Zheng, C. (2018). Role of groundwater in the dryland ecohydrological system: A case study of the Heihe River Basin. *Journal of Geophysical Research: Atmospheres*, 123, 6760–6776. <https://doi.org/10.1029/2018JD028432>
- Zhu, J. (2008). Equivalent parallel and perpendicular unsaturated hydraulic conductivities: Arithmetic mean or harmonic mean? *Soil Science Society of America Journal*, 72, 1226–1233. <https://doi.org/10.2136/sssaj2007.0337>