

Matrix Diffusion Controls Mountain Hillslope Groundwater Ages and Inferred Storage Dynamics

by Nicholas E. Thiros^{1,2} , Erica R. Woodburn¹, W. Payton Gardner³, James P. Dennedy-Frank⁴, and Kenneth H. Williams^{1,5}

Abstract

Groundwater age distributions provide fundamental insights on coupled water and biogeochemical processes in mountain watersheds. Field-based studies have found mixtures of young and old-aged groundwater in mountain catchments underlain by bedrock; yet, the processes that dictate these groundwater age distributions are poorly understood. In this work, we use the coupled ParFlow-CLM integrated hydrologic and EcoSLIM particle tracking models to simulate groundwater age distributions on a lower montane hillslope in the East River Watershed, Colorado (USA). We develop a convolution-based approach to propagate fracture-matrix diffusion processes to the EcoSLIM advection-dominated age distributions. We compare observed ^3H and ^4He concentrations from two groundwater wells against model predictions that have varying advective transport times and matrix diffusion magnitudes. Based on a Monte Carlo analysis that considers uncertain matrix and fracture parameters, we find that matrix diffusion is needed to jointly predict ^3H and ^4He observations at both wells. The advection-dominated age distributions lack adequate mixing of young and old-aged water to capture the observed co-occurrence of ^3H and ^4He . The model scenario that best matches the ^3H , ^4He , and water level observations when considering both advective flowpath and matrix diffusion mixing processes has a dynamic bedrock groundwater reservoir that is susceptible to considerable storage losses during low-snow periods. This dynamic groundwater system amplifies the need to assimilate deeper bedrock groundwater into watershed hydro-biogeochemical predictions. This work further highlights the importance of considering matrix diffusion when interpreting environmental tracers in bedrock groundwater systems.

Introduction

In mountainous watersheds, groundwater is recently demonstrated to be an important component in streamflow generation, ecologic function, and solute

(Brooks et al. 2021; Ryken et al. 2022; Carroll et al. 2024; Xie et al. 2024). Nonetheless, groundwater storage and flux dynamics in high-elevation headwater systems often remain uncertain due to the sparsity of direct observations and the complex dependencies on climate, topography, ecology, and geology (Somers and McKenzie 2020; Singha and Navarre-Sitchler 2022; Van Tiel et al. 2024). Groundwater storage volume dynamics and the timescales over which groundwater can support surface water and evapotranspiration (ET) fluxes are poorly understood (Hale et al. 2016; Hayashi 2020; Kirchner et al. 2023; Warix et al. 2023; Clayton et al. 2024), yet are critical to better predict the integrated hydro-biogeochemical response to climate change and drought conditions.

Integrated hydrologic models (IHMs) are a powerful tool that can be used to better understand the water and energy cycles in mountainous regions (Foster and Maxwell 2019; Thornton et al. 2022; Xu et al. 2023). IHMs simulate the coupled surface and subsurface hydrologic processes, including transient changes in groundwater storage volumes and fluxes. These types of models have been used to investigate the hydrologic and groundwater storage response due to climate change

¹Earth and Environmental Sciences Area, Lawrence Berkeley National Laboratory, Berkeley, California, USA

²Corresponding author: Nicholas E. Thiros, U.S. Geological Survey, Utah Water Science Center, West Valley City, UT, USA. Email: nthiros@usgs.gov

³Department of Geosciences, University of Montana, Missoula, Montana, USA

⁴Marine and Environmental Sciences, Northeastern University, Boston, Massachusetts, USA

⁵Rocky Mountain Biological Laboratory, Gothic, Colorado, USA

Article impact statement: Predicted mountain groundwater storage declines are sensitive to hydrologic and particle tracking models that consider matrix diffusion processes or not.

Received October 2024, accepted February 2025.

© 2025 The Author(s). *Groundwater* published by Wiley Periodicals LLC on behalf of National Ground Water Association.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](#) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

doi: 10.1111/gwat.13475

(Maina et al. 2022; Carroll et al. 2024; Engdahl 2024), providing valuable insights on a key water resource reservoir that is otherwise challenging to constrain. Nonetheless, IHMs remain simplified representations of complex mountain systems and often rely on validations using streamflow and groundwater level observations that are sparsely distributed throughout the years (Foster and Maxwell 2019; Maina et al. 2020; Thornton et al. 2022). It has been shown that these hydrometric datasets cannot uniquely constrain subsurface parameter distributions (for instance, through model parameter calibration), nor the groundwater mass transport and recharge rate timescales (McDonnell and Beven 2014; Thiros et al. 2021). Generally, multiple models that have varying process-model conceptualizations and hydrogeologic parameter distributions can equally fit the hydrometric observational data, which ultimately leads to process understanding and hydrologic prediction uncertainties (Enemark et al. 2019; Thiros et al. 2022).

Groundwater environmental tracers are sensitive to water velocities and can augment the information content provided by hydrometric data, helping to reduce groundwater flow and transport model parameter and prediction uncertainties (Portniaguine and Solomon 1998; McDonnell and Beven 2014; Starn et al. 2014; Thiros et al. 2021). However, few studies have directly compared environmental tracer observations to predictions made by IHMs in complex mountainous systems (e.g., Carroll et al. 2020). Furthermore, hydrologic models are typically compared to groundwater apparent or mean ages, which can be problematic given uncertainties and biases when deriving groundwater age metrics from environmental tracer observations using classic lumped-parameter model approaches (McCallum et al. 2014; Thiros et al. 2021). Consequently, there have been recent suggestions to assimilate environmental tracer concentrations directly into process-based models, rather than relying on groundwater age (Bethke and Johnson 2008; Suckow 2014; Turnadge and Smerdon 2014).

In a recently developed IHM (ParFlow-CLM) and Lagrangian particle tracking (EcoSLIM) model of a steep mountain hillslope in the Central Rocky Mountains of Colorado, USA (Thiros et al. 2024), an uncalibrated 2D hillslope model generally reproduced observed groundwater levels and mean ages inferred from environmental tracers at mid-slope and toe-slope locations. Despite the general agreement between observed and modeled mean ages in these wells finished within the fractured shale bedrock, the model poorly captured the mixing between pre-modern (>70 years) and modern (<70 years) groundwater age fractions, as suggested from the joint occurrence of ^3H and terrigenic ^4He ($^4\text{He}_{\text{terr}}$) in groundwater samples. In general, the simulated age distributions lacked either the modern or pre-modern component, suggesting an underprediction of flowpath mixing. While mixtures of modern and pre-modern fractions have also been observed in other mountain systems underlain by fractured bedrock, the mechanism generating the coupled young- and old-water mixing generally remains

uncertain (Hale et al. 2016; White et al. 2019; Manning et al. 2021; Markovich et al. 2021; Lerback et al. 2022). Rather than rely on typical lumped-parameter models that assume groundwater age distributions a priori, process-based IHMs can be useful to directly interrogate the salient mechanisms dictating the environmental tracer and groundwater age distributions. For instance, IHMs can capture spatially distributed and transient recharge events, flowpath mixing, and storage changes, which are all important processes to predict groundwater age distributions (e.g., Jiang et al. 2010; McCallum et al. 2017), yet are rarely considered in classic lumped-parameter model age dating techniques.

Matrix diffusion has been shown to be a key process controlling environmental tracer and groundwater age distributions in fractured rock systems (Bethke and Johnson 2002; Cook et al. 2005; Gardner et al. 2016; Rajaram 2021). Nonetheless, matrix diffusion is rarely considered in hillslope and watershed-scale particle tracking models that are coupled to IHMs. The particle tracking models presented in Thiros et al. (2024) better capture the groundwater age distributions dictated by advection and mechanical dispersion, which likely underestimates true mixing in heterogeneous subsurfaces (Sanford et al. 2017; Underwood et al. 2018). To our knowledge, no study has investigated the impact of matrix diffusion on simulating both pre-modern and modern groundwater mixing using an IHM coupled to a transient particle tracking code on a steep mountain hillslope. Evaluating the role of matrix diffusion on simulated environmental tracer distributions is important to better identify the potentially covarying and dynamic processes that dictate water and solute fluxes on this hillslope and others like it worldwide and how fluxes will change with projected climate change and drought.

In this work, we investigate the role of matrix diffusion in jointly predicting the co-occurrence of modern and pre-modern groundwater on a steep mountain hillslope. In particular, we assess the lack of explicit matrix diffusion in the EcoSLIM particle tracking simulations used in Thiros et al. (2024) as a potential source of the discrepancies between observed and simulated ^3H and ^4He distributions. We evaluate whether explicitly adding matrix diffusion to simulated advection-dominated groundwater age distributions helps jointly explain the young and old tracers (^3H and ^4He , respectively) at mid-slope and toe-slope groundwater wells. Furthermore, we identify the groundwater storage dynamics and potential resilience to drought conditions for a system where mixing is dominated by matrix diffusion. This information is critical to further develop understanding and modeling techniques that link measured environmental tracers and groundwater age to watershed water quantity and quality.

Methods

Study Site

The East River Watershed is within the Elk Mountains near Crested Butte, Colorado (USA; Figure 1a).

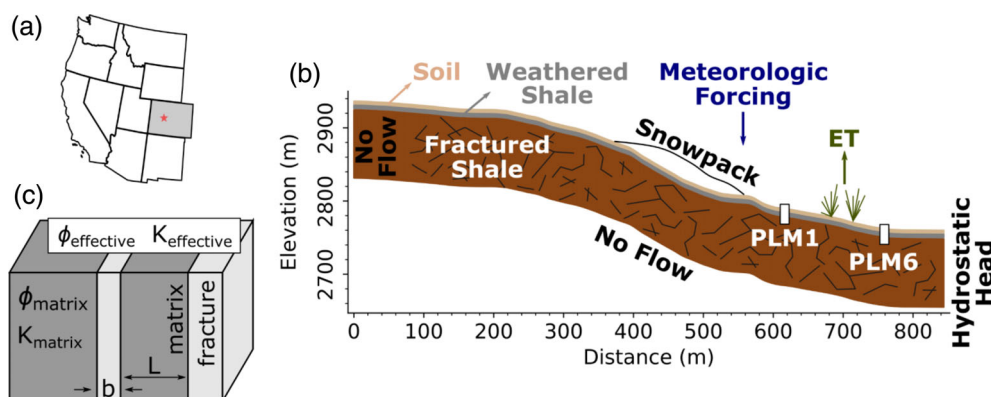


Figure 1. (a) Site location in the East River Watershed, Colorado, USA. (b) PLM hillslope ParFlow-CLM and EcoSLIM conceptual model and numerical domain. (c) Hypothetical fractured rock showing the parameters used to propagate the impact of matrix diffusion on simulated particle tracking age distributions.

The 300 km² watershed hosts broad research activities and is extensively monitored from bedrock to forest canopy. Details of the East River Watershed study area are described in Hubbard et al. (2018) and briefly introduced here. Elevations range from 2700 to 4100 m (mean 3266 m) and span river-valley montane to alpine eco-regions. Land cover generally consists of grasses and shrubs in the montane, evergreen and aspen forests in the sub-alpine, and bare soil and rocks in alpine regions. The climate is characterized as continental subarctic, with an average annual precipitation of 700 mm that predominantly occurs as winter snow (80% of total), daily mean air temperatures that range from −9.2°C to 9.8°C (annual mean 0°C; Carroll et al. 2018), and ET estimates from a lower-montane flux tower are 500 mm (Ryken et al. 2022). The East River hydrograph is dominated by snowmelt processes and timing. Flows peak in June and July, and baseflow occurs in the late fall and winter months. The East River in the study region is a meandering stream with an extensive alluvial floodplain, but steep incised bedrock channels exist in upstream regions of the watershed.

In this work, we focus on a single northeast-facing hillslope called the Pumphouse Lower Montane (PLM) hillslope (Figure 1b). The PLM hillslope is steep (average 22° degrees), 845 m long, and generally consists of ca. 1 m thick soil on top of weathered and fractured bedrock. The bedrock is marine-derived Cretaceous Mancos Shale that has experienced significant fracturing and faulting as a result of tectonic stress during the Laramide Orogeny, igneous intrusions, and glaciation and de-glaciation events (Miltenberger et al. 2021; Uhlemann et al. 2022; Quirk et al. 2024). Borehole geophysical surveys and core logs show fracturing occurs to depths of at least 80 m bls on the PLM hillslope; however, there exists considerable variability in the fracture density and spacing between adjacent hillslopes, likely due to discrete fault features (Miltenberger et al. 2021). The hillslope is generally divided into four vertical zones: soil ~0–2 m below land surface (mbls), highly weathered saprolite (~2–5 mbls), pervasively fractured bedrock (~5–100 mbls) and largely intact bedrock (>100 mbls; Tokunaga et al. 2019).

This work uses observational data from two groundwater wells finished within the fractured shale bedrock (Figure 1b). PLM1 is more upslope and finished at 10 m bls, and PLM6 is at the toe of the hillslope and finished at 10 m bls. Both wells are screened over the bottom ~2 m. Our recent work using environmental tracer observations shows that the fractured bedrock at PLM1 and PLM6 hosts water that is a mixture of modern and pre-modern ages (Thiros et al. 2023).

ParFlow-CLM and EcoSLIM Process-Based Modeling

A coupled ParFlow-CLM integrated hydrologic and EcoSLIM particle tracking model was developed for the PLM hillslope. The model is extensively described in Thiros et al. (2024) and we provide an overview here. The 2D hillslope transect follows land surface topography, starts at the local surface water divide, and ends at the East River (Figure 1b). The model has a constant depth of 100 m, and the subsurface was simplified to three hydrogeologic layers, each with constant thickness along the hillslope transect. The top, intermediate, and deepest layers are a 1 m thick soil, a 3 m thick weathered shale (saprolite), and a 96 m thick fractured shale bedrock, respectively (Figure 1b). Numerical grid horizontal resolution is 1.5 m, and vertical resolution ranges from 0.25 to 10 m in the shallowest and deepest layers, respectively. The model is forced using the North American Land Data Assimilation System (NLDAS-2) meteorological dataset at hourly timesteps from water year 2000 (WY2000) through WY2021 (Mitchell et al. 2004). Other boundary conditions include no-flow under the surface water divide and a constant hydrostatic head at the river corresponding to the annual average East River water depth. ParFlow-CLM models were dynamically spun up for 10 years using the 21-year hourly averaged NLDAS-2 forcing conditions.

EcoSLIM is a Lagrangian particle tracking model that uses ParFlow-CLM modeled infiltration, subsurface velocity, and ET fields to insert, advect, and remove particle mass from the subsurface (Maxwell et al. 2019). In EcoSLIM, particles enter the subsurface initialized

Table 1

ParFlow-CLM model parameters. *Top* and *bot* for the Exp Dec K model refer to the K_{fshale} at the shallowest and deepest fractured shale model layers, respectively. Log is the $\log_{10}$ of K , which has units of m s^{-1} .

Model	$\log K_{\text{soil}}$	$\log K_{\text{wshale}}$	$\log K_{\text{fshale}}$	$\alpha_{\text{soil}} (\text{m}^{-1})$	$\alpha_{\text{wshale}} (\text{m}^{-1})$	$n_{\text{soil}} (-)$	$n_{\text{wshale}} (-)$	$\phi_{\text{fshale}} (\%)$
A1	-3.6	-4.3	-7.1	4.4	4.1	4.8	3.2	10
A2	-3.4	-5.2	-7.7	0.4	1.5	2.1	3.5	10
A3	-4.3	-4.3	-7.0	3.6	2.3	3.4	2.8	10
A4	-3.8	-4.5	-6.9	3.4	3.3	3.1	3.3	10
A5	-3.8	-4.3	-6.5	2.2	3.6	4.4	1.8	10
Low ϕ	-3.6	-4.3	-7.1	4.4	4.1	4.8	3.2	5
High ϕ	-3.6	-4.3	-7.1	4.4	4.1	4.8	3.2	15
Exp Dec K	-3.6	-4.3	$-7.1^{\text{top}} - 8.1^{\text{bot}}$	4.4	4.1	4.8	3.2	10

with zero age at all timesteps and spatial coordinates with positive infiltration fluxes. Particles move as a function of the simulated velocity fields via the advection-dispersion equation until exiting the subsurface either as exfiltration to surface water or ET. Particle masses directly scale with ParFlow-CLM infiltration, ET, and exfiltration fluxes, allowing for the calculation of spatially and transiently distributed flux-weighted age distributions (Dennedy-Frank et al. 2024). For direct model-observation comparison, we use simulated resident groundwater age distributions at the discrete model cell locations and times that correspond to the sampling wells and dates. The EcoSLIM model was initialized with 5 particles with zero age at all model cells. A spin-up run recursively applied the last year of the transient ParFlow-CLM spinup runs (described above) for 3050 years (see Thiros et al. 2024 for details). EcoSLIM was then transiently run from WY2000 to WY2021 at daily timesteps.

In Thiros et al. (2024), we performed a Monte Carlo analysis on the ParFlow-CLM model considering adjustable saturated hydraulic conductivity (K) and van-Genuchten soil retention parameters α and n in the fractured shale (fshale), weathered shale (wshale) and soil model layers. The parameter list for the Monte Carlo analysis included: K_{soil} , K_{wshale} , K_{fshale} , α_{soil} , α_{wshale} , n_{soil} , and n_{wshale} . From the 128 Monte Carlo runs we selected the top five models that best matched observed groundwater level time series at PLM1 and PLM6 based on a minimum sum of squared error metric. These models are denoted as A1-A5, with A1 having the lowest sum of squared error. We additionally tested three variants of the A1 model that included a low and high porosity fractured shale (Low ϕ and High ϕ , respectively) and an exponentially decreasing K_{fshale} with depth (factor of 10 decrease per 100 m; Exp Dec K). Hydrogeologic parameters for all tested ParFlow-CLM models are given in Table 1. While this was not meant to be a formal calibration process, all of the tested models generally match the observed water level dynamics at PLM1 and PLM6 (see Thiros et al. 2024).

Observed and Simulated Environmental Tracers

Environmental tracer observations were sampled from the PLM1 and PLM6 wells in May 2021

(Thiros et al. 2023). In this work, we use the measured ^3H and $^4\text{He}_{\text{ter}}$ concentrations, as they, respectively, provide information on the young (decades) and old (millennia) portions of the groundwater age distribution. In contrast to Thiros et al. (2024), where EcoSLIM modeled and observed mean groundwater ages were qualitatively compared against each other, this work directly compares observed ^3H and $^4\text{He}_{\text{ter}}$ with modeled concentrations. Previous work has shown that directly comparing numerical model predictions with tracer concentrations can help limit the propagation of biases inherent in assuming groundwater age distributions and consequent mean ages (Suckow 2014; Thiros et al. 2021).

EcoSLIM simulations produce flux-weighted age distributions resulting from advection and dispersion that cannot be directly compared to measured environmental tracer concentrations. In order to calculate the environmental tracer concentration from the simulated advective-dispersive age distribution, we use the convolution integral (Cook and Herczeg 2000). The convolution integral:

$$C_{\text{out}}(t) = \int_0^\infty C_{\text{in}}(t-t') g(t') e^{-\lambda t'} dt' \quad (1)$$

relates the simulated flux-weighted age distribution from EcoSLIM $g(t')$ that is represented as a non-parametric histogram with the tracer-specific input signal $C_{\text{in}}(t-t')$ to produce flux-weighted tracer concentrations $C_{\text{out}}(t)$ at time t . In Equation 1, t' represents groundwater age (y) and $e^{-\lambda t'}$ accounts for first-order decay reactions with decay rate λ ($\lambda_{^3\text{H}} = 0.056 \text{ y}^{-1}$). The ^3H input series was taken from Michel et al. (2018) and $^4\text{He}_{\text{ter}}$ was assumed negligible in infiltrating water. $^4\text{He}_{\text{ter}}$ was assumed to accumulate in groundwater at a constant rate of $3.3 \times 10^{-11} (\text{cm}^3 \text{ STP g}^{-1} \text{ y}^{-1})$; see Thiros et al. (2023), equation 8 for details).

The EcoSLIM age distributions capture particle transport that includes advection and hydrodynamic dispersion subject to velocity distributions from boundary conditions and subsurface transport parameter distributions. The ParFlow-CLM and EcoSLIM modeling simulate velocities using effective porous media parameters and thus do not properly account for matrix diffusion processes in

fractured bedrock systems (Sanford et al. 2017). While discrete fracture network and dual-porosity models have been shown to capture the influence of matrix diffusion on age distributions in fractured rock media, they are generally computationally expensive and highly parameterized. In this work, we simulate the impact of matrix diffusion on the EcoSLIM advective-dispersive age distributions using a retention factor approach demonstrated in Painter et al. (2008). The convolution integral (Equation 1) is augmented to include particle retention due to diffusion into the matrix following Gardner et al. (2016):

$$C_{\text{out}}(t) = \int_0^\infty C_{\text{in}}(t - t_{\text{tran}}) \int_0^{t_{\text{tran}}} g_{\text{ret}}(t_{\text{tran}} - t'|\beta) \times \left(\frac{t'}{b}\right) g(t') e^{-\lambda t'} dt' dt_{\text{tran}}, \quad (2)$$

where t_{tran} is the total transport time (y), t' is the advective travel time (y), b is the effective fracture aperture (m), and β is a velocity-dependent transport resistance parameter given by $\beta = t'/b$. The matrix diffusion retention age distribution g_{ret} is a function of the retention time $t_{\text{ret}} = t_{\text{tran}} - t'$ (y) and assumes infinite matrix diffusion for a constant and uniform fracture aperture system:

$$g_{\text{ret}}(t_{\text{ret}}|\beta) = \frac{\kappa\beta}{2\sqrt{\pi}t_{\text{ret}}^{3/2}} \exp\left[-\frac{\kappa^2\beta^2}{4t_{\text{ret}}}\right], \quad (3)$$

where $\kappa = \phi_{\text{matrix}}\sqrt{\omega \times D_{\text{matrix}}}$ depends on the matrix porosity ϕ_{matrix} (–), matrix tortuosity ω (–), and diffusion coefficient D_{matrix} ($\text{m}^2 \text{s}^{-1}$). In this work $\omega = 0.1$ and $D_{\text{matrix}} = 2.3 \times 10^{-9} \text{m}^2 \text{s}^{-1}$ are considered constant and known and $g(t')$ is the flux-weighted advective age distribution from the EcoSLIM simulation.

We perform a Monte Carlo analysis over the uncertain b and ϕ_{matrix} parameters needed to parameterize Equation 2 and 3. Realizations of b were generated with the ‘cubic law’ for effective hydraulic conductivity K_{eff} (m s^{-1}) in fractured rock (Snow 1969):

$$b = [12L\mu/\rho g (K_{\text{eff}} - K_{\text{matrix}})]^{1/3}, \quad (4)$$

where L is the fracture spacing (m), μ is the dynamic viscosity of water ($\text{g m}^{-1} \text{s}^{-1}$), ρ is the density of water (g m^{-3}), g is the gravitational constant (m s^{-2}), and K_{matrix} is the hydraulic conductivity of the matrix block (m s^{-1}). Parameter prior distributions are given in Table 2 and

generally derived from borehole cores and geophysical NMR logs from a nearby well ($\sim 20 \text{m}$) on the PLM hillslope (ϕ_{matrix} , K_{matrix} , and L ; Uhlemann et al. 2022) and well pumping tests and prior model calibrations (K_{eff} ; Tokunaga et al. 2019). The prior parameter ranges are meant to be broad enough to capture the inherent uncertainty in these values, particularly due to the sparse observations available. A total of 1000 parameter-value sets were sampled using the quasi-random Sobol sequence technique (Sobol 1998) and ^3H and $^4\text{He}_{\text{terr}}$ predictions were recorded for each.

Results

Matrix Diffusion Controls on Environmental Tracer Concentrations

Figure 2 shows the joint ^3H and $^4\text{He}_{\text{terr}}$ predictions from Equation 1 considering the EcoSLIM advective-dispersive age distributions alone (open squares) and the advective-dispersive with matrix diffusion models (colored circles, Equation 2). Increasing mean age is a direct proxy for a higher magnitude of matrix diffusion, which occurs with low b or high ϕ_{matrix} (Figure S1). In other words, considering the matrix diffusion model (regardless of parameters values) increases the mean age relative to the advection-dominated EcoSLIM model (the youngest mean ages in Figure 2). Increasing the magnitude of matrix diffusion leads to higher $^4\text{He}_{\text{terr}}$ and generally lower ^3H concentrations. However, the general correlation between ^3H and mean age does not always hold due to the non-linear ^3H versus time input function (Figure S2), for example in Figure 2 (A4) at young mean ages. For multiple of the models that are characterized with EcoSLIM advection mean ages that exceed ~ 80 years (Figure 2: A2, Exp Dec K, and High ϕ), the simulated ^3H remains zero and is not impacted by matrix diffusion. $^4\text{He}_{\text{terr}}$ concentrations range from ~ 0 to $> 15 \text{cm}^3 \text{STPg}^{-1}$, which correspond to mean ages from 25 to $> 10,000$ years, respectively.

The quantitative match between observed and simulated ^3H and $^4\text{He}_{\text{terr}}$ concentrations is shown in Figure 3 using the χ^2 metric, where lower χ^2 indicates better matches. Like Figure 2, mean age positively correlates with the amount of matrix diffusion due to increased particle retention times. The A4, A5, and Low ϕ EcoSLIM advection models can generally explain the PLM1 ^3H observation, while the A5 and Low ϕ

Table 2
Parameter prior distributions used in the Monte Carlo analysis to estimate matrix diffusion magnitudes (Equations 3 and 4).

	K_{eff} (m s^{-1})	K_{matrix} (m s^{-1})	L (m)	ϕ_{matrix} (%)	b (m)
High bound	10^{-4}	10^{-9}	1.0	20	4.6×10^{-4}
Low bound	10^{-8}	10^{-12}	0.1	0.1	1.1×10^{-5}
Distribution	Log uniform	Log uniform	Log uniform	Log uniform	Equation 4

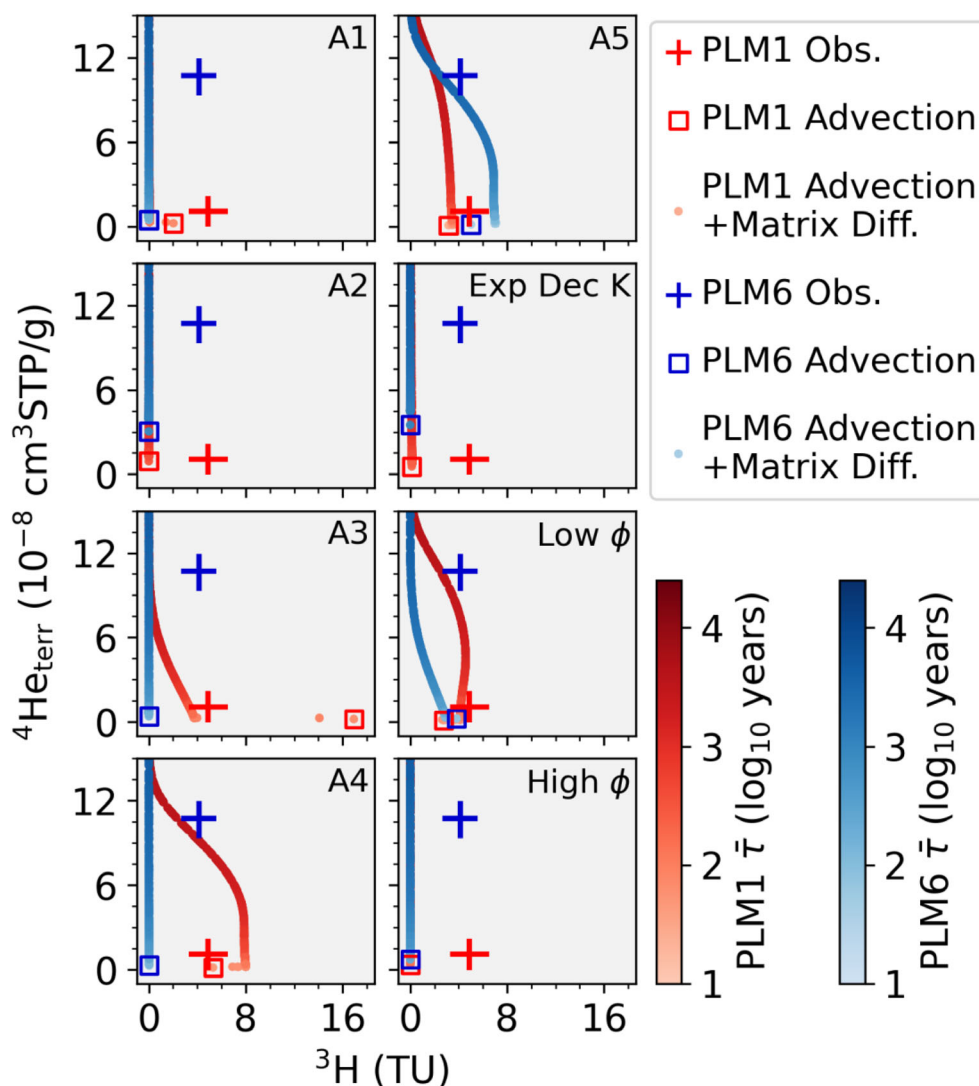


Figure 2. Simulated $^4\text{He}_{\text{terr}}$ and ^3H concentrations using the EcoSLIM model. Open squares are simulations that consider advection only and closed circles add the influence matrix diffusion based on the 1000-run Monte Carlo analysis. The mean age (τ) of the simulations is positively correlated with the magnitude of matrix diffusion. The observed environmental tracer concentrations are shown as colored error bars sizes corresponding to the marginal posterior probabilities calculated in Thiros et al. (2023).

advection models match the PLM6 ^3H observation. All of the advective models can predict the low $^4\text{He}_{\text{terr}}$ concentrations at PLM1. Alternatively, only the A5 advective model matches the observed $^4\text{He}_{\text{terr}}$ at PLM6. Taken together, the A4, A5 and Low ϕ models can generally predict both ^3H and $^4\text{He}_{\text{terr}}$ at PLM1 and none of the advection-dominated models accurately match both tracers at PLM6. In general, the EcoSLIM advection models underpredict ^3H at both wells and $^4\text{He}_{\text{terr}}$ at PLM6, suggesting the water is too young. The addition of matrix diffusion reduces the χ^2 and improves predictions of observed ^3H and $^4\text{He}_{\text{terr}}$ for the A3, A5, and Low ϕ models at PLM1. Similarly, matrix diffusion leads to improved predictions for both tracers at PLM6 for the A5 model. Comparing all the models shows that only the A5 model can jointly predict the ^3H and $^4\text{He}_{\text{terr}}$ observations at both PLM1 and PLM6 wells (Figures 2 and 3).

Figure 4a shows the b and ϕ_{matrix} parameters for the A5 model that best matches the observed ^3H and $^4\text{He}_{\text{terr}}$ (lowest χ^2 in Figure 3) and the subsequent 9 models with lowest χ^2 (open-gray circles in Figure 3). Parameters b and ϕ_{matrix} that predict the observations are positively correlated at both PLM1 and PLM6. This is due to the relationship between simulated age, b , and ϕ_{matrix} parameters in Equations 2 and 3, where age increases with ϕ_{matrix} and decreases with b (Figure S1). The age distributions for the A5 EcoSLIM advective model and the A5 matrix diffusion models that best match the observations are shown in Figure 4b and 4c. The A5 model has the youngest advective mean ages (~ 25 years) and largest K_{fshale} relative to the other models (Table 1). To jointly fit the observed ^3H and the low $^4\text{He}_{\text{terr}}$ concentrations at PLM1, the A5 model requires modest amounts of matrix diffusion that increase the median age

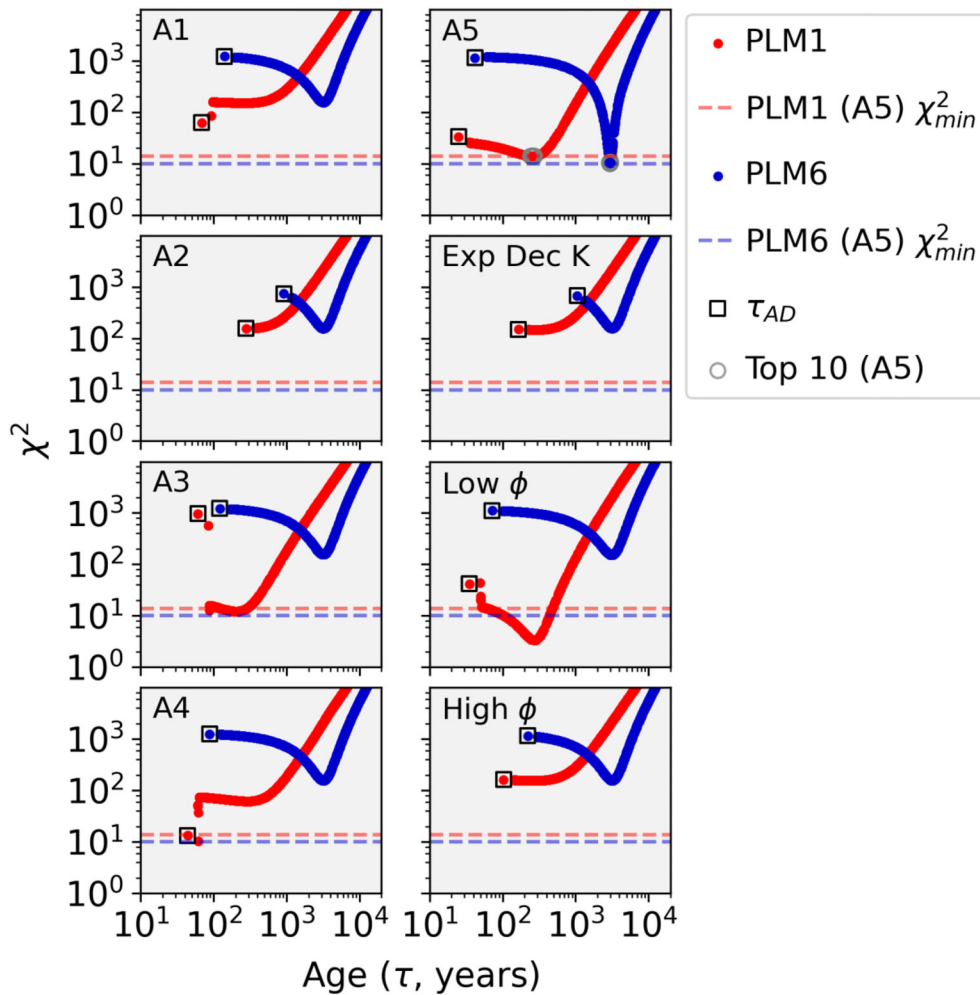


Figure 3. Simulated groundwater mean age (τ) versus chi-squared error metric (χ^2) that considers both ^3H and $^4\text{He}_{\text{terr}}$ observations for the advective-only EcoSLIM model (smallest mean age) and 1000 Monte Carlo runs that consider matrix diffusion. The horizontal-dashed lines correspond to the minimum χ^2 (best-fit model) achieved for the A5 model. The open-black square is the advection-dispersion only model and the open-gray circles are the 10 matrix-diffusion models with lowest χ^2 for the A5 model.

from 25 years for the advection simulation to 28 years when matrix diffusion is considered (open-square and closed-circle, respectively in Figure 4b). Alternatively, more matrix diffusion is needed to match the high $^4\text{He}_{\text{terr}}$ observation at PLM6, such that the age distribution has a long tail and the median age is near 400 years (Figure 4c). Analysis of the best-fit parameters in Figure 4a suggests the increase in matrix diffusion at PLM6 relative to PLM1 is due to higher ϕ_{matrix} .

Groundwater Storage Dynamics

While all models generally match the observed water level dynamics at PLM1 and PLM6 (see Thiros et al. 2024), only the A5 model jointly matches both environmental tracer observations at both wells. However, PLM1 and PLM6 are located near the bottom of the hillslope, and there are no direct groundwater level observations to constrain the model in the more upslope positions owing to access issues, slope steepness, and an inability to install monitoring wells. Figure 5a shows the

simulated groundwater levels at all hillslope positions and times (shaded regions). In the upslope positions (0–300 m), the groundwater levels are 25–35 m deeper for the A5 model compared to the other models. At the lower hillslope positions (>400 m) where the observation wells are located, all the models converge to approximately equivalent groundwater levels with similar temporal variations (shaded regions in Figure 5a). The overall deeper groundwater levels for the A5 model lead to lower hillslope groundwater storage volume magnitudes (Figure S3).

To better understand the temporal trends in the groundwater storage dynamics, Figure 5c shows the groundwater storage percent anomalies ($100 \times [s_i - s_{mu}]/s_{mu}$) that correspond to the daily storage volume (s_i) normalized by the 21-year (WY2000–WY2021) mean storage volume (s_{mu}). The A5 model groundwater storage anomaly has higher sensitivity compared to the other models on the multi-annual scale. In particular, the A5 model consistently has more positive and negative

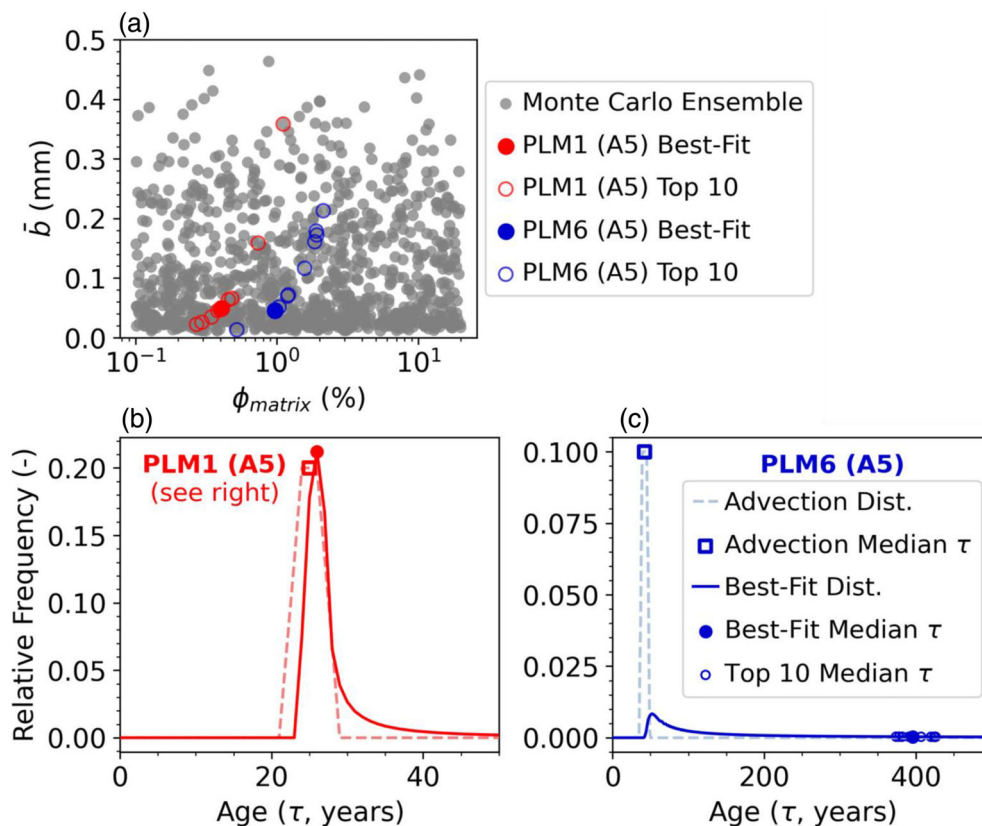


Figure 4. (a) Gray dots are model parameters ($N = 1000$) used in the Monte Carlo analysis to propagate the influence of matrix diffusion to the EcoSLIM advective age distribution for the A5 model. The parameter set that best jointly matches ^3H and ^4He at the PLM1 and PLM6 is indicated as a closed circle, while the next 9 parameter sets with the lowest sum-squared error are open circles. (b and c) Age distributions for the matrix diffusion model that best matches ^3H and ^4He jointly at PLM1 and PLM6 wells (solid lines) and the advective-only age distribution (dashed line).

storage anomalies during the wetter WY2009-WY2012 and drier WY2017-WY2021 periods, respectively (Figure 5b). The storage anomalies for all of the models generally match from WY2012 to WY2018. Following the low-snow WY2018 winter, the A5 model shows a more pronounced decline in groundwater storage relative to the other models. Notably, the declining storage trend in the A5 model propagates through WY2019, despite being characterized as a near-median precipitation year.

To further investigate the impact of multi-annual low-snow conditions on groundwater storage dynamics, we continue the ParFlow-CLM model with hypothetical forcing conditions. In particular, we force the A1 and A5 models for an extra 10 years past WY2021 using repeated forcing conditions that correspond to WY2016-WY2021 (left-most shaded region Figure 6). This time window represents the most pronounced storage losses in Figure 5a and 5c and a hypothetical future with prolonged snow-drought conditions, which is projected to occur in this region (Siirila-Woodburn et al. 2021). Although these extended hypothetical considerations do not consider other system changes that may occur in a watershed altered by a changing climate (e.g., vegetation reassembly or mortality), they are useful to quantify isolated impacts in a controlled manner. Comparing the groundwater storage anomalies during the extended simulation period

(Figure 6a) shows the A5 model continually loses storage (-2% relative to mean) while the A1 model is less impacted (-0.5% relative to mean). The storage anomaly decrease is the largest during the first 5-year drought period from WY2016 to WY2021, then has minor declines during the WY2026 to WY2031 future prediction periods. Figure 6b and 6c show the modeled annual cumulative runoff generation and ET from the entire hillslope. The A5 model generally has higher runoff and ET fluxes compared to the A1 model. The exception is for WY2017 (WY2022 and WY2027 for the future prediction scenarios), which had a large snowpack leading to relatively lower runoff and ET for the A5 model compared to the other years.

Discussion

Groundwater age distributions provide fundamental insights on how water moves through hillslopes and catchments (Sprenger et al. 2019). Our previous work using a lumped parameter age dating model shows the shallow groundwater on the steep PLM mountain hillslope is characterized as a mixture of modern and pre-modern ages that are on the order of thousands of years (Thiros et al. 2023). Lumped parameter models are considered a black-box approach that generally cannot uniquely identify the processes that control the

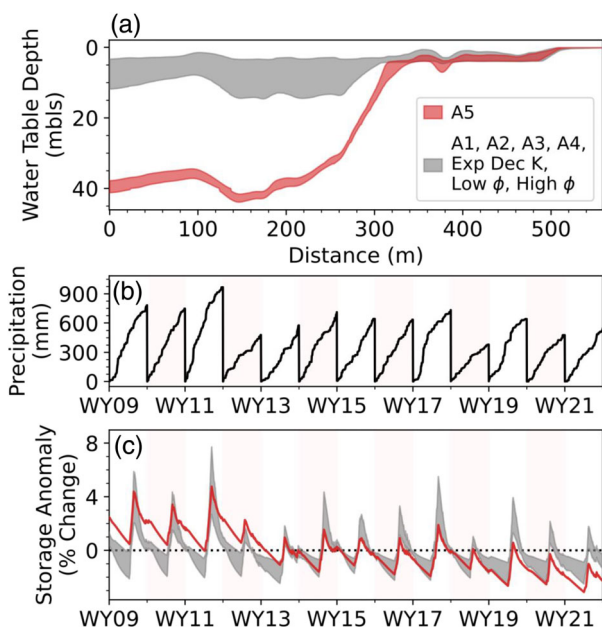


Figure 5. (a) Simulated water table depths at all timesteps from WY2017-WY2021 (shown as shaded region) and all hillslope transect positions. The gray regions correspond to all tested models other than the A5 model, which is shown in red. (b) The annual accumulated (water year) precipitation from the NLDAS-2 product. (c) Time series of groundwater storage anomalies from the 20-year mean storage calculated for each model individually. The shaded gray region corresponds to the range from all models shown in (a).

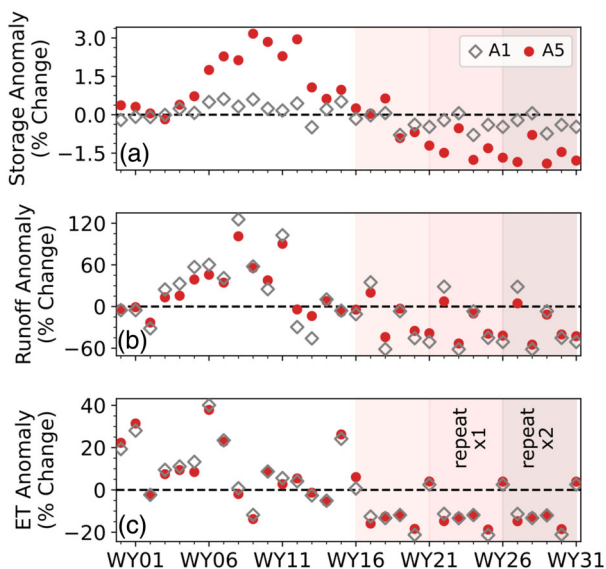


Figure 6. (a) Groundwater storage anomalies calculated on October 1 of the respective water year. The red shaded regions from WY2021-WY2026 and WY2026-WY2031 recursively apply the forcing conditions from WY2016-WY2021. (b) Simulated water year cumulative hillslope runoff anomalies and (c) ET water years cumulative anomalies.

mixing patterns, especially in complex mountain systems. Groundwater that is characterized as a mixture of young and old age components has been explained through multiple processes (e.g., Underwood et al. 2018), including flowpath mixing due to subsurface heterogeneity (Weissmann et al. 2002; Gardner et al. 2015), transience in recharge events (McCallum et al. 2017), and coupled fracture flow and matrix diffusion (Cook et al. 2005; Gardner et al. 2016; Sanford et al. 2017).

We find that the A5 model that invokes matrix diffusion to simulate the old water fraction best matches the observed ^3H and $^4\text{He}_{\text{terr}}$. This model has considerably different storage anomaly dynamics compared to the other models (Figure 5). The A5 model K_{fshale} is the highest among the models, leading to higher hydraulic diffusivity rates and younger resident mean groundwater ages at PLM1 and PLM6. The higher K_{fshale} results in a bedrock groundwater system that can fill and drain at faster rates compared to the other models tested, correspond to the more typical conceptual model of a low permeability bedrock mountain system (Condon et al. 2020). Similar to previous work that shows how differences in low versus high bedrock K can influence mountain groundwater storage dynamics (Gleeson and Manning 2008; Carroll et al. 2020; Gardner et al. 2020), the A5 model has less overall groundwater storage volume (Figure S3) and loses groundwater storage at a faster rate relative to the long-term mean during time periods characterized by low snow (Figure 5c). This model result is supported by the observed ~ 1 m decline in groundwater water levels in the past 7 years at PLM1 (Figure S4). Better understanding the potential ramifications of declining groundwater storage on streamflow generation, ET dynamics, and solute exports is critical for the management of mountain water resources (Wan et al. 2021; Ryken et al. 2022; Carroll et al. 2024).

Both advective flow and matrix diffusion are key processes that are needed to accurately predict the joint ^3H and $^4\text{He}_{\text{terr}}$ distribution in this system (Figure 2). The lower K_{wshale} to K_{fshale} ratios for the A5 model (relative to the other models) promote less lateral flow through the soil and saprolite layers and increased recharge into the deeper bedrock (Gardner et al. 2020). This process leads to higher flowpath velocities and younger advective resident ages in the fractured bedrock at PLM1 and PLM6, which supply the necessary ^3H to match the observations. Alternatively, matrix diffusion provides the old-aged water with elevated $^4\text{He}_{\text{terr}}$. Our results show that matrix diffusion is particularly important at the toe slope PLM6 well that has considerable $^4\text{He}_{\text{terr}}$, which cannot be explained by any of the advection-dominated model scenarios (Figure 3b). While high $^4\text{He}_{\text{terr}}$ can also be explained by other mechanisms, for instance, upwelling of deeper, old water (e.g., Bethke et al. 1999), these processes are generally captured within the modeled advective age distributions. For instance, the advective mean groundwater age at the toe slope PLM6 is older compared to the mid slope PLM1 due to modeled upwelling of deep flowpaths (Thiros et al. 2024). This upwelling process

has been found in other mountain catchment modeling studies (e.g., Kolbe et al. 2020), but cannot alone explain the observed age distribution in the East River Watershed at the hillslope scale. However, our model is limited to a 2D hillslope and will truncate regional flowpaths that can contain considerable $^4\text{He}_{\text{terr}}$ and age. Porosity distributions also influence simulated age and tracer concentrations. For instance, the assumed model porosities are within the higher range of values found within Uhlemann et al. (2022), potentially leading to water ages that are biased old. Further work is needed to test advective flowpath versus matrix diffusion mixing in controlling the old-aged component using 3D watershed models.

The best-fit fracture and matrix diffusion parameters for the A5 model at PLM1 and PLM6 are not equivalent (Figure 4a), highlighting variable water mixing processes between mid slope and toe slope wells. This modeling result is supported by the observed fracture distributions along the hillslope. PLM1 is in a mapped fault zone, with electrical resistivity tomography (ERT) anomalies suggesting that the fracture density is higher compared to the toeslope PLM6 well (Miltenberger et al. 2021). Within the modeling scenarios, the higher matrix diffusion at PLM6 compared to PLM1 is primarily due to higher ϕ_{matrix} , while b spans similar ranges. However, b and ϕ_{matrix} are correlated and further observations or regularization is needed to disaggregate the linked parameters. Possible key sources of information to better constrain the joint parameter distribution include analysis of fracture and porosity distributions from bedrock cores and/or estimation of spatial and vertical variations in bulk rock porosity derived from borehole geophysical measurements (e.g., neutron logging). Nonetheless, this work shows tracer observations that span disparate portions of the unknown age distribution can be valuable to better constrain the fracture and matrix parameters. For instance, multiple of the tested models can predict the observed ^3H at both wells and the $^4\text{He}_{\text{terr}}$ is needed to further falsify the models.

To our knowledge, this is one of the first works that interprets environmental tracer concentrations in a mountain system with a fully transient IHM. The ParFlow-CLM and EcoSLIM models do not require a priori assumptions on the shape of the advective age distribution, potentially reducing uncertainty inherent in classic groundwater age dating techniques that use lumped parameter models (Suckow 2014; Thiros et al. 2021). We find that interpreting multiple environmental tracers that span a broad range of groundwater ages is crucial to constrain the mixing processes in this system. For instance, multiple tested models can simulate the observed ^3H concentrations without matrix diffusion (Figure 2). The additional constraints on the age distribution provided by $^4\text{He}_{\text{terr}}$ suggest mixing with older-aged water, likely as a result of matrix diffusion processes. Our results further suggest that assimilating groundwater mean ages inferred from environmental tracers into the numerical model reduces the information content compared to assimilating the concentrations directly. For instance,

Thiros et al. (2024) showed that multiple hillslope models presented in this work can jointly explain the observed mean ages at PLM1 and PLM6 simultaneously. However, these models did not accurately capture the age distribution based on the co-occurrence of both ^3H and $^4\text{He}_{\text{terr}}$. Here, we show that matching both ^3H and $^4\text{He}_{\text{terr}}$ concentrations directly falsifies the EcoSLIM advective-only models and leads to a conceptual model with considerably different storage dynamics.

Our results suggest failure to consider matrix diffusion when assimilating environmental tracers into advection-dominated numerical models can lead to overestimation of the total drainable groundwater storage volumes. We show that high $^4\text{He}_{\text{terr}}$ concentrations that correspond to old groundwater ages can largely be the product of matrix diffusion, rather than large drainable hillslope storage volumes. This result has important implications for groundwater's ability to buffer streamflow declines during drought (Carroll et al. 2024; Xie et al. 2024). Groundwater in low K matrix blocks is considered part of the immobile storage and is likely unavailable to significantly contribute to the mobile storage that is connected to streamflow on the annual timescales (Birkel et al. 2011). This result is supported by recent work that shows matrix diffusion controls the tail of watershed transit time distributions (Rajaram 2021). Our results also have important implications for how environmental tracer and groundwater age information is assimilated into numerical models. Simulating old groundwater ages is often achieved by increasing model depths such that there are higher storage volumes and increased contributions from deeper flowpaths. In contrast, our results show that such techniques can lead to overestimation of the total drainable storage volumes, which can lead to inaccurate predictions of groundwater's ability to buffer streamflow during drought conditions.

Despite the agreement between simulated and observed groundwater levels and environmental tracer concentrations, there are multiple sources of uncertainty in the modeling setups. Model non-commensurability and equifinality likely leads to estimated parameters that do not represent the point scale observations available from wells and boreholes (Beven 2006). For instance, parameter equifinality and high porosity estimates can lead to the A5 model that has higher K_{fshale} compared to the other models better matching the observed tracer distributions. Similarly, the low ϕ_{matrix} estimated in Figure 4a relative to Manger (1963) and Uhlemann et al. (2022) suggests model structural errors are influencing the parameter inference (e.g., Thiros et al. 2021). A potential model error is the simplified hydrostatic head boundary condition applied beneath the river. This static boundary condition does not include no-flow conditions at depths below the river, potentially leading to an underprediction of the upwelling of deeper and older-aged groundwater into the river, floodplain, and PLM6 areas. The applied boundary condition represents the best feasible approximation that is readily implemented within the ParFlow-CLM modeling framework, nonetheless,

contributes to the uncertainty in simulated groundwater ages. As a result of these structural model errors, forecasts and predictions with these best-fit parameters will contain significant uncertainty. While model uncertainty analysis techniques that characterize full posterior distributions (Thiros et al. 2022) and use multiple, differing conceptual models (Enemark et al. 2019) would improve the reliability of estimated parameters and predictions, such techniques remain computationally infeasible with current IHMs. A further limitation of this work is the lack of water level data in the upper hillslope regions that can be used to falsify models. Such data is challenging to acquire in steep mountain systems that are difficult, and in many instances, impossible to access.

Conclusion

We use coupled, integrated hydrologic, particle tracking, and convolution integral models to investigate the influence of matrix diffusion on groundwater environmental tracers (^3H and $^4\text{He}_{\text{terr}}$) concentrations on a steep, lower montane hillslope in Central Colorado, USA. We find that matrix diffusion is a key process needed to predict the mixing of premodern and modern groundwater ages in this mountainous system. The model scenario that best matches the ^3H , $^4\text{He}_{\text{terr}}$, and water level observations when considering both advective flowpath and matrix diffusion mixing processes has a dynamic bedrock groundwater reservoir that is susceptible to considerable storage losses during low-snow periods. The conceptual model of a dynamic bedrock groundwater system follows observational (Manning and Caine 2007) and high-fidelity modeling studies (e.g., Gardner et al. 2020; Carroll et al. 2024), calling into question the often assumed depth of the critical zone being the top of the bedrock (see Condon et al. 2020 for discussion). We show the value of constraining parameters and predictions from an IHM using both groundwater level and multiple environmental tracer concentrations that span a broad distribution of ages. Our work emphasizes the need for in-situ bedrock groundwater observations to better constrain watershed-scale hydro-biogeochemical dynamics in mountain systems (Singha and Navarre-Sitchler 2022).

Acknowledgment

This material is based upon work supported as part of the Watershed Function Scientific Focus Area funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Contract No. DE-AC02-05CH11231. Further support is provided in part from the Department of Energy (DOE) EPSCoR grant DE918 FOA-0002215.

Data Availability Statement

ParFlow-CLM and EcoSLIM model input decks and field observation data used in this work are available in Thiros et al. (2024).

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article. Supporting Information is generally *not* peer reviewed.

Figure S1. Relationship between mean age, fracture aperture, and matrix porosity.

Figure S2. Time series of ^3H infiltration concentrations.

Figure S3. Modeled absolute storage dynamics for the A1–A5 models.

Figure S4. Observed groundwater levels at PLM1.

References

- Bethke, C.M., and T.M. Johnson. 2008. Groundwater age and groundwater age dating. *Annual Review of Earth and Planetary Sciences* 36, no. 1: 121–152.
- Bethke, C.M., and T.M. Johnson. 2002. Paradox of groundwater age. *Geology* 30, no. 4: 385–388.
- Bethke, C.M., X. Zhao, and T. Torgersen. 1999. Groundwater flow and the ^4He distribution in the Great Artesian Basin of Australia. *Journal of Geophysical Research: Solid Earth* 104, no. B6: 12999–13011.
- Beven, K. 2006. A manifesto for the equifinality thesis. *Journal of Hydrology* 320, no. 1–2: 18–36.
- Birkel, C., C. Soulsby, and D. Tetzlaff. 2011. Modelling catchment-scale water storage dynamics: Reconciling dynamic storage with tracer-inferred passive storage. *Hydrological Processes* 25, no. 25: 3924–3936.
- Brooks, P.D., A. Gelderloos, M.A. Wolf, L.R. Jamison, C. Strong, D.K. Solomon, G.J. Bowen, S. Burian, X. Tai, S. Arens, L. Briefer, T. Kirkham, and J. Stewart. 2021. Groundwater-mediated memory of past climate controls water yield in snowmelt-dominated catchments. *Water Resources Research* 57, no. 10: e2021WR030605.
- Carroll, R.W.H., R.G. Niswonger, C. Ulrich, C. Varadharajan, E.R. Siirila-Woodburn, and K.H. Williams. 2024. Declining groundwater storage expected to amplify mountain stream-flow reductions in a warmer world. *Nature Water* 2, no. 5: 419–433.
- Carroll, R.W.H., A.H. Manning, R. Niswonger, D. Marchetti, and K.H. Williams. 2020. Baseflow age distributions and depth of active groundwater flow in a snow-dominated mountain headwater basin. *Water Resources Research* 56: e2020WR028161.
- Carroll, R.W.H., L.A. Bearup, W. Brown, W. Dong, M. Bill, and K.H. Williams. 2018. Factors controlling seasonal groundwater and solute flux from snow-dominated basins. *Hydrological Processes* 32, no. 14: 2187–2202.
- Clayton, N., E. Knappe, A.M. White, H.R. Martens, D.F. Argus, N. Lau, A.A. Borsa, R. Bendick, and W.P. Gardner. 2024. Elastic deformation as a tool to investigate watershed storage connectivity. *Communications Earth & Environment* 5, no. 1: 110.
- Condon, L.E., K.H. Markovich, C.A. Kelleher, J.J. McDonnell, G. Ferguson, and J.C. McIntosh. 2020. Where is the bottom of a watershed? *Water Resources Research* 56, no. 3: e2019WR026010. <https://doi.org/10.1029/2019WR026010>
- Cook, P.G., A.J. Love, N.I. Robinson, and C.T. Simmons. 2005. Groundwater ages in fractured rock aquifers. *Journal of Hydrology* 308, no. 1–4: 284–301.
- Cook, P.G., and A.L. Herczeg. 2000. *Environmental Tracers in Subsurface Hydrology*. Boston, MA, USA: Springer Science & Business Media.
- Dennedy-Frank, P.J., A. Visser, F.Z. Maina, and E.R. Siirila-Woodburn. 2024. Investigating mountain watershed

- headwater-to-groundwater connections, water sources, and storage selection behavior with dynamic-flux particle tracking. *Journal of Advances in Modeling Earth Systems* 16, no. 8: e2023MS003976.
- Enemark, T., L.J.M. Peeters, D. Mallants, and O. Batelaan. 2019. Hydrogeological conceptual model building and testing: A review. *Journal of Hydrology* 569: 310–329.
- Engdahl, N.B. 2024. Impacts of permeability uncertainty in a coupled surface-subsurface flow model under perturbed recharge scenarios. *Water Resources Research* 60, no. 3: e2023WR035975.
- Foster, L.M., and R.M. Maxwell. 2019. Sensitivity analysis of hydraulic conductivity and Manning's n parameters lead to new method to scale effective hydraulic conductivity across model resolutions. *Hydrological Processes* 33, no. 3: 332–349.
- Gardner, W.P., K. Jensco, Z.H. Hoylman, R. Livesay, and M.P. Maneta. 2020. A numerical investigation of bedrock groundwater recharge and exfiltration on soil mantled hillslopes. *Hydrological Processes* 34, no. 15: 3311–3330.
- Gardner, W.P., M. Hokr, H. Shao, A. Balvin, H. Kunz, and Y. Wang. 2016. Investigating the age distribution of fracture discharge using multiple environmental tracers, Bedrichov Tunnel, Czech Republic. *Environmental Earth Sciences* 75, no. 20: 1–16.
- Gardner, W.P., G. Hammond, and P. Lichtner. 2015. High performance simulation of environmental tracers in heterogeneous domains. *Groundwater* 53, no. S1: 71–80.
- Gleeson, T., and A.H. Manning. 2008. Regional groundwater flow in mountainous terrain: Three-dimensional simulations of topographic and hydrogeologic controls. *Water Resources Research* 44, no. 10: 1–16.
- Hale, C.V., J.J. McDonnell, M.K. Stewart, D.K. Solomon, J. Doolittle, G.G. Ice, and R.T. Pack. 2016. Effect of bedrock permeability on stream base flow mean transit time scaling relationships: 2. Process study of storage and release. *Water Resources Research* 52: 1375–1397.
- Hayashi, M. 2020. Alpine hydrogeology: The critical role of groundwater in sourcing the headwaters of the world. *Groundwater* 58, no. 4: 498–510.
- Hubbard, S.S., K.H. Williams, D. Agarwal, J. Banfield, H. Beller, N. Bouskill, E. Brodie, R. Carroll, B. Dafflon, D. Dwivedi, N. Falco, B. Faybishenko, R. Maxwell, P. Nico, C. Steefel, H. Steltzer, T. Tokunaga, P.A. Tran, H. Wainwright, and C. Varadharajan. 2018. The East River, Colorado, watershed: A mountainous community testbed for improving predictive understanding of multi-scale hydrological-biogeochemical dynamics. *Vadose Zone Journal* 17, no. 1: 180061.
- Jiang, X., L. Wan, M.B. Cardenas, S. Ge, and X. Wang. 2010. Simultaneous rejuvenation and aging of groundwater in basins due to depth-decaying hydraulic conductivity and porosity. *Geophysical Research Letters* 37, no. 5: 2010GL042387.
- Kirchner, J.W., P. Benettin, and I. Van Meerveld. 2023. Instructive surprises in the hydrological functioning of landscapes. *Annual Review of Earth and Planetary Sciences* 51, no. 1: 277–299.
- Kolbe, T., J. Marçais, J. De Dreuzy, T. Labasque, and K. Bishop. 2020. Lagged rejuvenation of groundwater indicates internal flow structures and hydrological connectivity. *Hydrological Processes* 34, no. 10: 2176–2189.
- Lerback, J.C., B.B. Bowen, C.E. Humphrey, D.P. Fernandez, J.A. Bernau, S.J. Macfarlan, E. Schniter, and J.J. Garcia. 2022. Geochemistry and provenance of springs in a Baja California Sur mountain catchment. *Groundwater* 60, no. 2: 295–308. <https://doi.org/10.1111/gwat.13177>
- Maina, F.Z., E.R. Siirila-Woodburn, and P.J. Denny-Frank. 2022. Assessing the impacts of hydrodynamic parameter uncertainties on simulated evapotranspiration in a mountainous watershed. *Journal of Hydrology* 608: 127620. <https://doi.org/10.1016/j.jhydrol.2022.127620>
- Maina, F.Z., E.R. Siirila-Woodburn, M. Newcomer, Z. Xu, and C. Steefel. 2020. Determining the impact of a severe dry to wet transition on watershed hydrodynamics in California, USA with an integrated hydrologic model. *Journal of Hydrology* 580: 124358.
- Manger, E. 1963. *Porosity and Bulk Density of Sedimentary Rocks*. United States Department of the Interior. 1–44.
- Manning, A.H., L.B. Ball, R.B. Wanty, and K.H. Williams. 2021. Direct observation of the depth of active groundwater circulation in an alpine watershed. *Water Resources Research* 57, no. 2: 2020WR028548. <https://doi.org/10.1029/2020WR028548>
- Manning, A.H., and J.S. Caine. 2007. Groundwater noble gas, age, and temperature signatures in an alpine watershed: Valuable tools in conceptual model development. *Water Resources Research* 43, no. 4: 2006WR005349. <https://doi.org/10.1029/2006WR005349>
- Markovich, K.H., L.E. Condon, K.C. Carroll, R. Purtschert, and J.C. McIntosh. 2021. A mountain-front recharge component characterization approach combining groundwater age distributions, noble gas thermometry, and fluid and energy transport modeling. *Water Resources Research* 57, no. 1: 1–21.
- Maxwell, R.M., L.E. Condon, M. Danesh-Yazdi, and L.A. Bearup. 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, no. 1: 1–10.
- McCallum, J.L., P.G. Cook, S. Dogramaci, R. Purtschert, C.T. Simmons, and L. Burk. 2017. Identifying modern and historic recharge events from tracer-derived groundwater age distributions. *Water Resources Research* 53, no. 44: 1039–1056.
- McCallum, J.L., P.G. Cook, C.T. Simmons, and A.D. Werner. 2014. Bias of apparent tracer ages in heterogeneous environments. *Groundwater* 52, no. 2: 239–250.
- McDonnell, J.J., and K. Beven. 2014. Debates – The future of hydrological sciences: A (common) path forward? A call to action aimed at understanding velocities, celerities and residence time distributions of the headwater hydrograph. *Water Resources Research* 50, no. 6: 5342–5350.
- Michel, R.L., B.C. Jurgens, and M.B. Young. 2018. *Tritium Deposition in Precipitation in the United States, 1953–2012*. U.S. Geological Scientific Investigations Report 2018–5086.
- Miltenberger, A., S. Uhlemann, T. Mukerji, K. Williams, B. Dafflon, L. Wang, and H. Wainwright. 2021. Probabilistic evaluation of geoscientific hypotheses with geophysical data: Application to electrical resistivity imaging of a fractured bedrock zone. *Journal of Geophysical Research: Solid Earth* 126, no. 9: 1–20.
- Mitchell, K.E., et al. 2004. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *Journal of Geophysical Research: Atmospheres* 109, no. D7: 2003JD003823.
- Painter, S., V. Cvetkovic, J. Mancillas, and O. Pensado. 2008. Time domain particle tracking methods for simulating transport with retention and first-order transformation. *Water Resources Research* 44, no. 1: 1–11.
- Portniaguine, O., and D.K. Solomon. 1998. Parameter estimation using groundwater age and head data, Cape Cod, Massachusetts. *Water Resources Research* 34, no. 4: 637.
- Quirk, B.J., I.J. Larsen, and A.J. Hidy. 2024. Latest Pleistocene glacial chronology and paleoclimate reconstruction for the East River watershed, Colorado, USA. *Quaternary Research* 119: 86–98.
- Rajaram, H. 2021. Matrix diffusion as a mechanism contributing to fractal stream chemistry and long-tailed transit time

- distributions. *Geophysical Research Letters* 48, no. 18: e2021GL094292.
- Ryken, A.C., D. Gochis, and R.M. Maxwell. 2022. Unraveling groundwater contributions to evapotranspiration and constraining water fluxes in a high-elevation catchment. *Hydrological Processes* 36, no. 1: 1–14.
- Sanford, W.E., L.N. Plummer, G. Casile, E. Busenberg, D.L. Nelms, and P. Schlosser. 2017. Using dual-domain advective-transport simulation to reconcile multiple-tracer ages and estimate dual-porosity transport parameters. *Water Resources Research* 53, no. 6: 5002–5016.
- Siirila-Woodburn, E.R., A.M. Rhoades, B.J. Hatchett, L.S. Huning, J. Szinai, C. Tague, P.S. Nico, D.R. Feldman, A.D. Jones, W.D. Collins, and L. Kaatz. 2021. A low-to-no snow future and its impacts on water resources in the western United States. *Nature Reviews Earth and Environment* 2, no. 11: 800–819.
- Singha, K., and A. Navarre-Sitchler. 2022. The importance of groundwater in critical zone science. *Groundwater* 60, no. 1: 27–34.
- Snow, D.T. 1969. Anisotropic permeability of fractured media. *Water Resources Research* 5, no. 6: 1273–1289.
- Sobol, I.M. 1998. On Quasi-Monte Carlo integrations. *Mathematics and Computers in Simulation* 47, no. 2–5: 103–112.
- Somers, L.D., and J.M. McKenzie. 2020. A review of groundwater in high mountain environments. *Wiley Interdisciplinary Reviews: Water* 7, no. 6: e1475.
- Sprenger, M., C. Stumpp, M. Weiler, W. Aeschbach, S.T. Allen, P. Benettin, M. Dubbert, A. Hartmann, M. Hrachowitz, J.W. Kirchner, J.J. McDonnell, N. Orłowski, D. Penna, S. Pfahl, M. Rinderer, N. Rodriguez, M. Schmidt, and C. Werner. 2019. The demographics of water: A review of water ages in the critical zone. *Reviews of Geophysics* 57, no. 3: 800–834.
- Starn, J.J., C.T. Green, S.R. Hinkle, A.C. Bagtzoglou, and B.J. Stolp. 2014. Simulating water-quality trends in public-supply wells in transient flow systems. *Ground Water* 52: 53–62.
- Suckow, A. 2014. The age of groundwater – Definitions, models and why we do not need this term. *Applied Geochemistry* 50: 222–230.
- Thiros, N.E., E.R. Siirila-Woodburn, M. Sprenger, K.H. Williams, J.P. Denny-Frank, R.W.H. Carroll, and W.P. Gardner. 2024. Old-aged groundwater contributes to mountain hillslope hydrologic dynamics. *Journal of Hydrology* 635: 131193.
- Thiros, N.E., E.R. Siirila-Woodburn, P.J. Denny-Frank, K.H. Williams, and W.P. Gardner. 2023. Constraining bedrock groundwater residence times in a mountain system with environmental tracer observations and Bayesian uncertainty quantification. *Water Resources Research* 59, no. 2: e2022WR033282.
- Thiros, N.E., W.P. Gardner, M.P. Maneta, and D.J. Brinkerhoff. 2022. Quantifying subsurface parameter and transport uncertainty using surrogate modelling and environmental tracers. *Hydrological Processes* 36, no. 11: e14743.
- Thiros, N.E., W.P. Gardner, and K.L. Kuhlman. 2021. Utilizing environmental tracers to reduce groundwater flow and transport model parameter uncertainties. *Water Resources Research* 57, no. 7: e2020WR028235. <https://doi.org/10.1029/2020WR028235>
- Thornton, J.M., R. Therrien, G. Mariéthoz, N. Linde, and P. Brunner. 2022. Simulating fully-integrated hydrological dynamics in complex Alpine headwaters: Potential and challenges. *Water Resources Research* 58: e2020WR029390.
- Tokunaga, T.K., J. Wan, K.H. Williams, W. Brown, A. Henderson, Y. Kim, A.P. Tran, M.E. Conrad, M. Bill, R.W.H. Carroll, W. Dong, Z. Xu, A. Lavy, B. Gilbert, S. Carrero, J.N. Christensen, B. Faybishenko, B. Arora, E.R. Siirila-Woodburn, R. Versteeg, J.H. Raberg, J.E. Peterson, and S.S. Hubbard. 2019. Depth- and time-resolved distributions of snowmelt-driven hillslope subsurface flow and transport and their contributions to surface waters. *Water Resources Research* 55, no. 11: 9474–9499.
- Turnadge, C., and B.D. Smerdon. 2014. A review of methods for modelling environmental tracers in groundwater: Advantages of tracer concentration simulation. *Journal of Hydrology* 519: 3674–3689.
- Uhlemann, S., B. Dafflon, H.M. Wainwright, K.H. Williams, B. Minsley, K. Zamudio, B. Carr, N. Falco, C. Ulrich, and S. Hubbard. 2022. Surface parameters and bedrock properties covary across a mountainous watershed: Insights from machine learning and geophysics. *Science Advances* 8: eabj2479.
- Underwood, S.C., J.L. McCallum, and P.G. Cook. 2018. Physical and chemical controls on the simultaneous occurrence of young and old groundwater inferred from multiple age tracers. 54, no. 11: 9514–9532.
- Van Tiel, M., et al. 2024. Cryosphere–groundwater connectivity is a missing link in the mountain water cycle. *Nature Water* 2, no. 7: 624–637.
- Wan, J., T.K. Tokunaga, W. Brown, A.W. Newman, W. Dong, M. Bill, C.A. Beutler, A.N. Henderson, N. Harvey-Costello, M.E. Conrad, N.J. Bouskill, S.S. Hubbard, and K.H. Williams. 2021. Bedrock weathering contributes to subsurface reactive nitrogen and nitrous oxide emissions. *Nature Geoscience* 14, no. 4: 217–224.
- Warix, S.R., A. Navarre-Sitchler, A.H. Manning, and K. Singha. 2023. Local topography and streambed hydraulic conductivity influence riparian groundwater age and groundwater-surface water connection. *Water Resources Research* 59, no. 9: e2023WR035044.
- Weissmann, G.S., Y. Zhang, E.M. LaBolle, and G.E. Fogg. 2002. Dispersion of groundwater age in an alluvial aquifer system. *Water Resources Research* 38, no. 10: 13–16.
- White, A., B. Moravec, J. McIntosh, Y. Olshansky, B. Paras, R.A. Sanchez, T.P.A. Ferré, T. Meixner, and J. Chorover. 2019. Distinct stores and the routing of water in the deep critical zone of a snow-dominated volcanic catchment. *Hydrology and Earth System Sciences* 23, no. 11: 4661–4683.
- Xie, J., X. Liu, S. Jasechko, W.R. Berghuijs, K. Wang, C. Liu, M. Reichstein, M. Jung, and S. Koirala. 2024. Majority of global river flow sustained by groundwater. *Nature Geoscience* 17, no. 8: 770–777.
- Xu, Z., E.R. Siirila-Woodburn, A.M. Rhoades, and D. Feldman. 2023. Sensitivities of subgrid-scale physics schemes, meteorological forcing, and topographic radiation in atmosphere-through-bedrock integrated process models: A case study in the Upper Colorado River basin. *Hydrology and Earth System Sciences* 27, no. 9: 1771–1789.