Water Resources Research

RESEARCH ARTICLE
10.1029/2018WR023863

Key Points:
- Modeled subsurface flow and transport in mountain watersheds are strongly affected by missing low-order streams/ridges in numerical meshes
- Baseflow is biased lower/higher for low-/high-order streams, while mean age and solute concentration are biased higher for all stream orders
- The magnitude of the bias is influenced by the degree of subsurface heterogeneity and reaction weathering rates for weathering products

Supporting Information:
- Supporting Information S1

Correspondence to:
C. Wang,
chao.wang@student.nmt.edu

Citation:
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*. https://doi.org/10.1029/2018WR023863

Received 8 AUG 2018
Accepted 5 DEC 2018
Accepted article online 14 DEC 2018

The Importance of Capturing Topographic Features for Modeling Groundwater Flow and Transport in Mountainous Watersheds

Chao Wang[1], Jesus D. Gomez-Velez[1,2], and John L. Wilson[1]

[1]Hydrology Program, Department of Earth and Environmental Science, New Mexico Institute of Mining and Technology, Socorro, NM, USA, [2]Department of Civil and Environmental Engineering, Vanderbilt University, Nashville, TN, USA

Abstract Physics-based distributed hydrological models that include groundwater are widely used to understand and predict physical and biogeochemical processes within watersheds. Typically, due to computational limitations, watershed modelers minimize the number of elements used in domain discretization, smoothing or even ignoring critical topographic features. We use an idealized model to investigate the implications of mesh refinement along streams and ridges for modeling three-dimensional groundwater flow and transport in mountainous watersheds. For varying degrees of topographic complexity level (TCL), which increases with the level of mesh refinement, and geological heterogeneity, we estimate and compare steady state baseflow discharge, mean age, and concentration of subsurface weathering products. Results show that ignoring lower-order streams or ridges diminishes flow through local flow paths and biases higher the contribution of intermediate and regional flow paths, and biases baseflow older. The magnitude of the bias increases for systems where permeability rapidly decreases with depth and is dominated by shallow flow paths. Based on a simple geochemical model, the concentration of weathering products is less sensitive to the TCL, partially due to the thermodynamic constraints on chemical reactions. Our idealized model also reproduces the observed emergent scaling relationship between the groundwater contribution to streamflow and drainage area, and finds that this scaling relationship is not sensitive to mesh TCL. The bias effects have important implications for the use of hydrological models in the interpretation of environmental tracer data and the prediction of biogeochemical evolution of stream water in mountainous watersheds.

Plain Language Summary Topography controls the movement of water and solutes within watersheds and their export to streams via groundwater drainage. Mainly due to computational limitations, it is common practice to use numerical meshes that capture the high-order (i.e., large) streams and ignore or undersample low-order (i.e., small) streams and their ridges. However, low-order streams are hotspots for groundwater drainage and solute export, and their ridges recharge a significant amount of water feeding aquifers and alluvial valleys downstream. By systematically comparing groundwater flow and transport characteristics from models with different degrees of topographic fidelity along streams and ridges, we find that even though ignoring low-order streams and ridges has a modest effect in the net amount of discharge generated by the whole watershed, it causes significant changes in the complex network of subsurface flow paths, which distribute water and solutes throughout the system. These changes bias estimates of recharge and discharge, residence times, and solute fluxes to streams, and they are more evident in watersheds with tight bedrock and primarily shallow groundwater flow. This bias hampers our ability to model and predict hydrologic response and directly influences the management of water resources and water quality for human consumption and ecosystem functioning.

1. Introduction

Physics-based distributed hydrological models are designed to explicitly represent watershed flow and transport processes while assimilating a broad range of physical and chemical observations (e.g., Houser et al., 1998; Liu et al., 2012; Park & Xu, 2009). These models can reproduce complex spatiotemporal variations of storage and fluxes and are instrumental in the quest for deep, mechanistic understanding of hydrologic systems (Beven, 1989, 2001; Clark et al., 2011). In particular, these models have been widely used to quantify

runoff generation and solute transport in the environment (e.g., Bao et al., 2017; Li & Sivapalan, 2011; Li et al., 2017; Partington et al., 2013; Sudicky et al., 2008), interpret physical and chemical observations (e.g., Kollet & Maxwell, 2008; Maxwell et al., 2016; Su et al., 2016), and predict hydrological and biogeochemical evolution of water under present and future climate and land use conditions (e.g., Goderniaux et al., 2015; Green & Wang, 2008; Hartmann et al., 2017).

Even though distributed hydrologic models have the potential to explain and predict flow and transport in natural environments, they are prone to uncertainty (Beven, 2001; Gupta et al., 2012), which imposes limitations in our ability to accurately reproduce both integrated (or net) and distributed watershed responses. Extensive research has focused on understanding the uncertainties for three key stages of model development (Gupta et al., 2012; Gupta & Nearing, 2014): (i) conceptual representation, (ii) system parametrization, and (iii) numerical interpolation and integration. In this work, we focus on the *numerical interpolation and integration* step and, in particular, the importance of capturing key topographic features with the numerical meshes used to integrate and model subsurface flow and transport characteristics. As part of this effort, we also explore how changes in geological heterogeneity, an important component of the *system parametrization* step, affect our findings. Quantifying the implications of selective mesh refinement can improve our understanding of the uncertainties involved in modeling and ultimately play a crucial role in our ability to make predictions and decisions of economic and social importance.

Topography exerts a dominant control in the movement of water and solutes through surface and subsurface pathways (Beven & Freer, 2001; Beven & Kirkby, 1979; Cardenas, 2007; Frei et al., 2010; Li et al., 2013; O'Loughlin, 1986; Riml & Wörman, 2011; Stocker et al., 2014; Tóth, 1963; Wilcox et al., 2007; Wolock et al., 1990). Previous work has shown that using coarse representations of the landscape can result in significant modeling uncertainty (e.g., Singh et al., 2015; Sulis et al., 2011; Vieux, 1993; Vivoni et al., 2005; Zhang & Montgomery, 1994). However, given the fractal nature of landscapes, and the limitations in observations and computational and numerical capabilities, this uncertainty cannot be removed by indiscriminately refining the numerical mesh. This raises the question of *which features should be captured and what are the implications of neglecting some of them?*

Previous studies have investigated the importance of accurately representing topography or domain boundaries for the simulation of flow and solute transport in hydrologic systems at different scales and from different aspects. For example, in modeling channel flow, Horritt et al. (2006) find that finer meshes can better represent the channel boundary and drive hydraulic features like recirculation zones. For hyporheic flow modeling, poor representation of bathymetry leads to underestimation of small-scale and overestimation of large-scale hyporheic exchange fluxes (Chow et al., 2018). In groundwater recharge estimation, Stoertz and Bradbury (1989) used a fully saturated groundwater flow model to show that total modeled recharge increases with decreasing mesh resolution, which is explained by the model's ability to capture local flow paths as the mesh size is reduced. In watershed rainfall-runoff modeling, selectively retaining slope, curvature and topographic index can influence simulated runoff generation mechanisms and hydrograph shape (Mahmood & Vivoni, 2011; Vivoni et al., 2004, 2005). A coarse mesh can greatly decrease the spatial variability of landscape curvature and the spatial variability of simulated soil moisture content during wet periods (Kuo et al., 1999). On the other hand, coarse representations of the topography can increase river discharge volumes as well as water table depth and soil water storage (Sulis et al., 2011), and in particular, mesh coarsening within hydrologically significant near-stream regions can shift hydrograph response from slow subsurface runoff to quick surface runoff (Vivoni et al., 2005). Over large scales, analyses of land-surface models have shown that explicit representation of stream features can improve simulated bidirectional land-to-river hydrological interactions, which can further affect inundation extent and concomitant biogeochemical and ecological processes (Shen et al., 2016).

Notice that most of the previous studies exploring the importance of topographic discretization have focused on relatively shallow processes, mainly taking place within the critical zone. Little attention has been paid to the implications of the complex and nested system of groundwater flow paths characterizing hydrologic systems (Tóth, 1963; Wörman et al., 2007). In this regard, topography strongly influences these connections from local to regional scales. At the local scale, topographic variations in stream beds drive hyporheic exchange (e.g., Boano et al., 2014), which at the same time is modulated by large-scale subsurface fluxes (Buffington & Tonina, 2009; Caruso et al., 2016; Tonina & Buffington, 2009). At the intermediate scale, landscape variables like drainage area, hillslope length, and stream bank gradient can influence the amount and age of baseflow

and control stream chemistry composition (Singh et al., 2016). And finally, at the regional scale, topographic characteristics like ridge-valley asymmetry and stream incision depth can affect the amount of interwatershed groundwater flow and mountain block recharge (Welch & Allen, 2012; Wilson & Guan, 2004) and explain the observed patterns of stream water quality and residence times (Cardenas, 2007; Kirchner et al., 2000; Kirchner & Neal, 2013; Wörman et al., 2007).

Of particular interest are ubiquitous landscape features such as stream networks and ridgelines, given their important role as spatial hotspots for discharge and recharge (Freeze & Witherspoon, 1967; Tóth, 1963; Winter, 1999; Winter et al., 2003). During and between rainfall events, variably saturated contributing areas expand and contract around river network, modulating the generation of runoff from the local to the catchment scale (Glaser et al., 2016; Güntner et al., 1999). Moreover, the spatial location (topology and geometry) of the river network strongly influences the drainage of groundwater into streams and the proportions of local, intermediate, and regional groundwater flow (Gleeson & Manning, 2008; Tóth, 1963). Similarly, the spatial distribution of ridges (location and relative elevation) controls the emergence of areas with enhanced recharge and the dynamics of the groundwater divide and interbasin groundwater exchange (Welch et al., 2012; Welch & Allen, 2012; Winter et al., 2003). Given the importance of these features for subsurface flow systems, it stands to reason that capturing them in numerical models can have important implications for interpretation and prediction. However, typical hydrological modeling efforts at the scale of mountain watersheds only use high-resolution meshes along selected high-order streams (e.g., Ball et al., 2014; Chen, 2015; Glaser et al., 2016), ignoring the lower-order ones and typically neglecting the effect of ridges.

This paper studies the effect of capturing streams and ridges on modeling groundwater flow and transport in mountainous watersheds. We hypothesize that *capturing different amount of streams and ridges by the mesh can have a systematic effect on the simulated groundwater flow field, and further affects the spatial distribution of river baseflow and its mean age and solute concentrations*. To the best of our knowledge, this is the first time that a systematic sensitivity analysis of this kind has been performed.

To test our hypothesis, we use the topography of the Rio Hondo watershed in Northern New Mexico, USA. This is a typical mountainous watershed with steep and complex topography characterized by significant contributions of deep groundwater to streamflow that tend to increase with drainage area (Frisbee et al., 2017; Tolley, 2014; Tolley et al., 2015). The hydrological and geochemical trends observed in this watershed are consistent with the conceptualization of a three-dimensional, topography-driven groundwater flow system (Tóth, 1963, 2009) and a three-dimensional catchment-mixing watershed conceptual model (Frisbee et al., 2011). Thus, the Rio Hondo watershed provides an ideal test bed to explore our hypothesis with the aid of numerical experiments that allow us to generalize our conclusions.

Even though our test bed is a real watershed, we use an idealized steady state groundwater flow conceptualization where the water table is a replica of the topography, that is, the *Tothian assumption* (Tóth, 1963). This assumption has been widely used in the literature, and it has the advantage of being parsimonious — we avoid the need to solve a highly nonlinear dynamic boundary condition for the water table — while offering useful insight into the key features and controls for flow and transport in a myriad of hydrologic systems, including hyporheic zones and regional groundwater flow systems (Cardenas, 2008; Freeze & Witherspoon, 1967; Gomez & Wilson, 2013; Marklund & Wörman, 2011; Tóth, 2009, 1963; Wörman et al., 2007). In particular, the groundwater table is often found to closely follow land surface topography and therefore termed *topography-controlled*, in regions with high water table ratio as previously defined by Haitjema and Mitchell-Bruker (2005) and further explored at the continental scale by Gleeson et al. (2011). In a recent study, Condon and Maxwell (2015) explicitly calculated the relative importance of topographic head gradients and pressure head gradients in driving groundwater flow for the contiguous United States and found that areas where groundwater flow is primarily driven by topographic head gradients correspond well with the topography-controlled water table regions identified by Gleeson et al. (2011), a finding that further supports the use of the Tothian assumption.

Even though this simple conceptualization has limitations (Bresciani et al., 2016; Haitjema & Mitchell-Bruker, 2005), it has proven useful to explain key processes in natural systems (Cardenas, 2007; Freeze & Witherspoon, 1967; Frisbee et al., 2011, 2017; Genereux et al., 2013; Gomez & Wilson, 2013; Tóth, 2009; Wörman et al., 2007). For example, Tóth (1963) used this assumption to explain the nesting of flow paths in a two-dimensional, cross-sectional basin model. He found that topography factors such as basin's aspect ratio, regional slope, and local relief explain the development and relative importance of local, intermediate, and regional groundwater flow systems, consistent with physical and geochemical observations. Gleeson

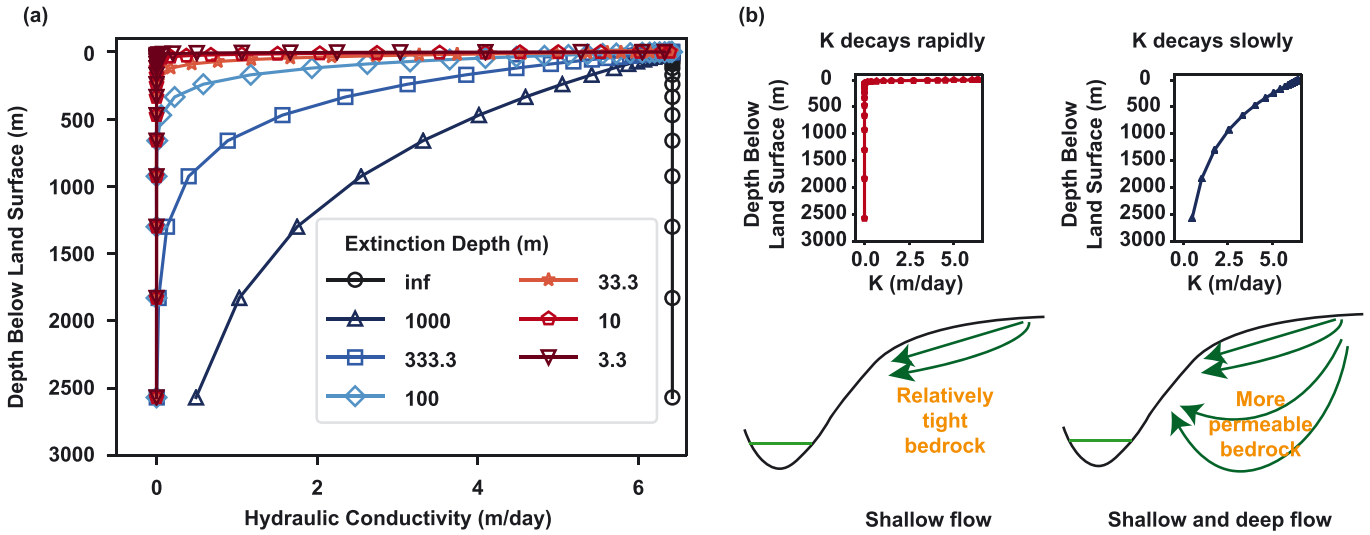

Figure 1. (a) One homogeneous and six exponentially decaying hydraulic conductivity scenarios representing different geological conditions. (b) Flow path distributions within two representative hillslope cross sections. For hillslopes where hydraulic conductivity decays rapidly with depth, shallow flow dominates. For hillslopes where hydraulic conductivity decays relatively slowly with depth, both shallow and deep groundwater flow are present in the system.

and Manning (2008) extended this analysis to three dimensions, simulating steady state groundwater flow in synthetic mountainous terrains—with recharge and seepage top boundary conditions. They found that, with a topography-controlled water table, the relative contribution of regional flow is well correlated with the mean elevation of the first-order stream and is sensitive to the topographic variables such as flank slope and amplitude of the relief (Gleeson & Manning, 2008). These modeling results provide the theoretical basis for distinguishing and mapping recharge-controlled and topography-controlled water table at the continental scale (Gleeson et al., 2011) and for conceptualizing groundwater systems based on topographic, climatic, and geologic data in regions with sparse groundwater observations.

In the spirit of these previous efforts, we develop model scenarios of different topographic complexity levels (TCLs) by progressively capturing more streams and ridges with mesh refinement. Homogeneous and exponentially decaying hydraulic conductivity fields are used to represent different geological heterogeneities (i.e., bedrock permeability conditions). The effect of capturing topographic features on groundwater circulation and the spatial patterns of baseflow and its mean age and solute concentrations are analyzed.

2. Methods

2.1. Model Configuration and Subsurface Heterogeneity

Our analyses use the topography of the Rio Hondo watershed (i.e., topographic template) with different geological heterogeneity scenarios (i.e., geologic template), effectively generating a series of distinct systems where we explore the effect of mesh resolution in detail. Our systems include one homogeneous and six heterogeneous hydraulic conductivity fields, resulting in a total number of seven. To this end, we use an exponentially decaying isotropic hydraulic conductivity conceptualization (Louis, 1972; Figure 1a):

$$K(\vec{x}) = K_0 \exp[-\alpha\, d(\vec{x})] \tag{1}$$

where K_0 is the hydraulic conductivity at the surface (L/T), $d(\vec{x})$ is depth below land surface (L) with $\vec{x} = [x, y, z]$ the coordinate vector (L) and coordinate z positive upward, and α is a decay rate (L^{-1}). The decay rate can be used to define the extinction depth $d_e = 1/\alpha$ (L), which represents the depth where the hydraulic conductivity is approximately 37% of the surface value (i.e., an e-fold decrease). For all simulations $K_0 = 6.43$ m/day, representative of sand, and $\alpha \in \{0.001, 0.003, 0.01, 0.03, 0.1, 0.3\}$ m^{-1} ($d_e \in \{1000, 333.3, 100, 33.3, 10, 3.3\}$ m), consistent with previous studies (Cardenas & Jiang, 2010; Ingebritsen & Manning, 1999; Saar & Manga, 2004). Equation (1) allows us to mimic systems that range from shallow, hillslope-dominated flow in the critical zone to deep flow in permeable fractured bedrock (Figure 1b).

2.1.1. Topographic Complexity Scenarios

We generate a series of three-dimensional meshes that capture topographic variations with different degrees of fidelity by maintaining a constant value for the largest element size and progressively increasing resolu-

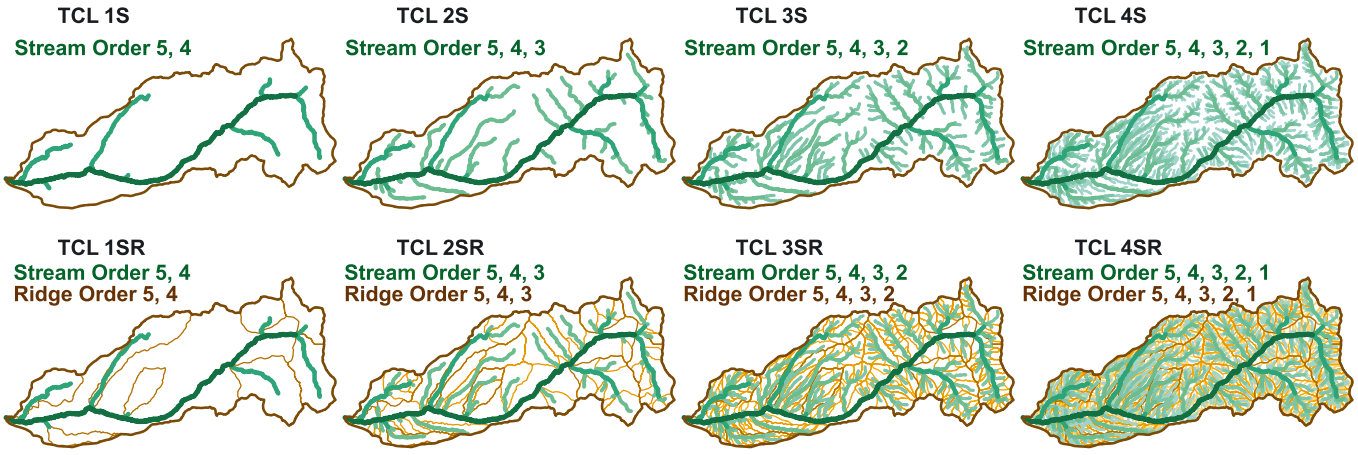

Figure 2. Streams (green lines) and ridges (brown polygons) used for mesh refinement, with lighter shades representing lower orders. Each topographic complexity level (TCL) refines the mesh along corresponding linear features. TCL4S captures all the streams and is used as an "ideal" base case to evaluate models with other meshes missing some streams (TCL1S, TCL2S, and TCL3S). TCL*icl*SR (*icl* = 1, 2, 3, 4) are used to evaluate the effect of missing ridges.

tion along selected streams and ridges—we refer to this process as increasing the TCL. To this end, a digital elevation model (DEM) for the Rio Hondo watershed, with a resolution of 1/3 arc sec, was obtained from the U.S. Geological Survey National Elevation Dataset. The DEM is hydrologically corrected (Tarboton et al., 1991) and streams and ridges of subwatersheds up to order of 5 are extracted using the Hydrology toolset implemented in Esri® ArcGIS®. The extracted stream network is then used as a template to generate meshes with varying TCLs.

First, we discretize the watershed into a two-dimensional (2-D) triangular mesh using Leapfrog Geo®. The maximum node spacing (80 m), constant for all the topographic complexity scenarios, is determined as half the average divide-to-stream distance of the watershed, which is approximated as half the reciprocal of drainage density (Dingman, 2015; Horton, 1932; Rodríguez-Iturbe & Rinaldo, 1997). Second, to capture local-scale topographic variability, we prescribe mesh nodes with a finer resolution (20 m) along the polylines that define the streams and ridges. Using this approach, we increase the TCL, by progressively including lower-order streams and their corresponding ridges (Figure 2). For example, TCL1S, the lowest complexity level, only refines the mesh along the fourth- and fifth-order streams and TCL1SR refines the same streams as TCL1S plus their corresponding ridges. TCL4SR, the highest complexity level, refines the mesh along all the streams and ridges from order of 1 to 5. Refinement features for the TCLs explored are shown in Figure 2.

Finaly, each triangular mesh is extruded vertically to create 3-D triangular prism elements. The top of the model is the land surface elevation. The bottom of the modeling domain is 3,000 m below and parallel to the land surface. The simulation domain is vertically divided into 29 parallel layers with a thickness that increases exponentially downward, ranging from 0.1 m for the top layer to 860 m for the bottom layer.

2.1.2. Groundwater Flow

For each geological heterogeneity scenario, groundwater flow is simulated with each of the TCL meshes. Given that our focus is on topography-driven groundwater flow and its implications for baseflow generation and solute transport, we simulate saturated, steady state groundwater flow with no sources/sinks (equation (2)) and assume that the water table is a replica of the land surface elevation (i.e., the terrain elevation corresponds to the prescribed head at the top boundary, equation (3); Freeze & Witherspoon, 1967; Haitjema & Mitchell-Bruker, 2005; Tóth, 1963). Lateral and bottom boundaries are assumed to be impervious (equation (4)). The mathematical statement under these assumptions is given by

$$\nabla \cdot [K \nabla h] = 0 \quad \text{in} \quad \Omega \tag{2}$$

$$h(x, y, z = z_0(x, y)) = z_0(x, y) \quad \text{on} \quad \partial\Omega_1 \tag{3}$$

$$\vec{n} \cdot [K \nabla h] = 0 \quad \text{on} \quad \partial\Omega_2 \text{and} \partial\Omega_3 \tag{4}$$

where h is hydraulic head (L), K is hydraulic conductivity (L/T; equation (1)), z_0 is land surface elevation (L) and is assumed as specified water table for the top boundary condition, $\vec{n}$ is the outward normal vector at the boundary, Ω is the modeling domain, $\partial\Omega_1$ is the top boundary, $\partial\Omega_2$ is the lateral boundary, and $\partial\Omega_3$ is the bottom boundary.

The mathematical statement (2)–(4) is solved with HydroGeoSphere (Therrien et al., 2010), a physically based, fully integrated surface-subsurface 3-D flow and transport code based on a control volume finite element method.

2.1.3. Groundwater Age

Similarly, for each geological heterogeneity scenario, baseflow mean age is simulated with each of the TCL meshes. The mathematical statement for the mean age, A (T), is given by

$$-\nabla \cdot \vec{q}A + \nabla \cdot \theta \mathbf{D} \nabla A + \theta = 0 \quad \text{in} \quad \Omega \tag{5}$$

$$A = 0 \quad \text{on} \quad \partial\Omega_{1,\text{in}} \tag{6}$$

$$\vec{n} \cdot (\theta \mathbf{D} \nabla A) = 0 \quad \text{on} \quad \partial\Omega_{1,\text{out}} \tag{7}$$

$$\vec{n} \cdot (\vec{q}A - \theta \mathbf{D} \nabla A) = 0 \quad \text{on} \quad \partial\Omega_2 \text{and} \, \partial\Omega_3 \tag{8}$$

where $\vec{q} = -K\nabla h$ is the Darcy flux vector (L/T), $\mathbf{D}$ is macrodispersion tensor (L^2/T), θ is porosity (-), and $\partial\Omega_{1,\text{in}}$ and $\partial\Omega_{1,\text{out}}$ represent the areas with flow entering and leaving the domain along the top boundary respectively. Mathematically these boundaries are defined as $\partial\Omega_{1,\text{in}} = \{\vec{x} | \vec{n} \cdot \vec{q} < 0\}$, $\partial\Omega_{1,\text{out}} = \{\vec{x} | \vec{n} \cdot \vec{q} > 0\}$, $\partial\Omega_{1,\text{in}} + \partial\Omega_{1,\text{out}} = \partial\Omega_1$. The macrodispersion tensor $\mathbf{D}$ is given by $\theta\mathbf{D} = (\alpha_l - \alpha_t)\frac{\vec{q}\cdot\vec{q}^T}{|\vec{q}|} + \alpha_t |\vec{q}| \mathbf{I}$ (Bear, 1972), with α_l and α_t the longitudinal and transverse dispersivities (L), respectively, and $\mathbf{I}$ is the identity tensor. We use $\theta = 0.375$, $\alpha_l = 100$ m, and $\alpha_t = 30$ m for all simulations, which are within the range of field observations (e.g., Gelhar et al., 1992). HydroGeoSphere is used to solve equations (5)–(8), and a detailed description of the theory and application of age distributions and their moments can be found in Goode (1996), Varni and Carrera (1998), Ginn (1999), and Gomez and Wilson (2013).

2.2. Metrics

2.2.1. Baseflow Into Streams

We use baseflow to evaluate the effect of TCL on flow patterns under various geological scenarios. Simulated water flux leaving the domain through the top boundary $\partial\Omega_1$ (i.e., groundwater discharge) is assigned to a stream node, using the topography and a steepest descent algorithm, and treated as baseflow, whether or not the mesh is locally refined. Total baseflow into a stream i of order ω for a given TCL mesh ($cl \in \{1S, 1SR, \ldots, 4S, 4SR\}$) is calculated as follows:

$$Q^{cl}_{b,\omega,i} = \sum_{j \in \Omega^{\omega,i}_{\text{node}}} |Q_j| \tag{9}$$

where Q_j is the volumetric flux associated with node j (L^3/T), $\Omega^{\omega,i}_{\text{node}}$ is the set of nodes with groundwater leaving the domain (i.e., node $j \in \Omega_{1,\text{out}}$) and draining into stream $l_{\omega,i}$, determined with the flow directions estimated during the DEM correction and stream network extraction. $i = 1, 2, \ldots, N_\omega$ with N_ω the number of streams of order ω ($\omega = 1, 2, \ldots, 5$ for the Rio Hondo watershed). The total flux discharging into streams and eventually leaving the watershed for the entire stream network is calculated as follows:

$$Q^{cl}_{b,T} = \sum_{\forall Q_j < 0} |Q_j| \tag{10}$$

We use expressions (9) and (10) to propose a normalized baseflow for each stream $l_{\omega,i}$:

$$\rho^{cl}_{b,\omega,i} = \frac{Q^{cl}_{b,\omega,i}}{Q^{cl}_{b,T}} \tag{11}$$

Finally, to quantify the effect of progressively refining the mesh along lower-order streams, we calculate the relative difference of the normalized baseflow between models of TCL*icl*S (*icl*=1,2,3) and the model with the mesh that captures all stream orders (TCL4S):

$$\Delta\rho^{icl\text{S}}_{b,\omega,i} = \frac{\rho^{icl\text{S}}_{b,\omega,i} - \rho^{4\text{S}}_{b,\omega,i}}{\rho^{4\text{S}}_{b,\omega,i}} \times 100\% \tag{12}$$

then, for example, a positive value of this metric indicates more baseflow into stream link i of order ω will result when a mesh with TCL*icl*S is selected.

Similarly, the effect of capturing ridges is evaluated by comparing results of meshes TCL*icl*S and meshes TCL*icl*SR (*icl*=1,2,3,4):

$$\Delta\rho^{icl\text{SR}}_{b,\omega,i} = \frac{\rho^{icl\text{S}}_{b,\omega,i} - \rho^{icl\text{SR}}_{b,\omega,i}}{\rho^{icl\text{SR}}_{b,\omega,i}} \times 100\% \tag{13}$$

2.2.2. Flushing Intensity

The flushing intensity (Gomez-Velez et al., 2014; Zlotnik et al., 2011), an integrated measure of the capacity of the flow system to transport water mass by advection at different depths, is used:

$$F(z) = \frac{1}{S}\int_S \sqrt{q_x^2(\vec{x}) + q_y^2(\vec{x}) + q_z^2(\vec{x})}\,\mathrm{d}s, \tag{14}$$

where S is the area of the layer over which the magnitude of flux is integrated. Flushing intensity provides a quantitative measure of the groundwater flow system structure at different depths. The flushing intensity of models with different TCLs is compared with the reference model under the same geologic condition. In particular, we define the relative differences

$$\Delta F^{icl\mathrm{S}} = \frac{F^{icl\mathrm{S}} - F^{4\mathrm{S}}}{F^{4\mathrm{S}}} \times 100\% \tag{15}$$

to quantify the effect of progressively refining the mesh along lower-order streams, and

$$\Delta F^{icl\mathrm{SR}} = \frac{F^{icl\mathrm{S}} - F^{icl\mathrm{SR}}}{F^{icl\mathrm{SR}}} \times 100\% \tag{16}$$

to quantify the effect of refining the mesh along ridges.

2.2.3. Baseflow Mean Age

Similar to the baseflow rate, we estimate the flux-weighted mean age of baseflow for each stream $I_{\omega,i}$ as follows:

$$A^{cl}_{b,\omega,i} = \frac{1}{Q^{cl}_{b,\omega,i}} \sum_{j\in\Omega^{\omega,i}_{\mathrm{node}}} |Q_j| A_j. \tag{17}$$

We use the same approach as equations (12) and (13) to quantify the relative difference in baseflow mean age for each stream. In this case, we use the notation $\Delta A^{icl\mathrm{S}}_{b,\omega,i}$ (%) and $\Delta A^{icl\mathrm{SR}}_{b,\omega,i}$ (%) for the case with streams only and streams and ridges, respectively.

2.2.4. Chemical Weathering Flux

The relative magnitude of groundwater residence time compared to the time required for mineral weathering to reach equilibrium controls the concentration of solutes from chemical weathering in catchments, influences chemical evolution of stream water, and influences CO_2 emission and thus Earth's temperature (Maher, 2011; Maher & Chamberlain, 2014). Using the analytical solution of a 1-D advection-reaction equation to describe the solute concentration along one particular flow path, and assuming an exponential distribution of groundwater residence times, Maher and Chamberlain (2014) propose a lumped model describing the mean solute concentration C (M/L^3) at the catchment outlet:

$$C = C_{\mathrm{eq}}\frac{\tau Da}{1+\tau Da} = C_{\mathrm{eq}}\left(1 - \frac{1}{1+\tau Da}\right), \tag{18}$$

where C_{eq} (M/L^3) is the thermodynamic limit of equilibrium concentration, $\tau = \exp(2) \approx 7.4$ is a scaling factor, which ensures the concentration from one particular flow path reaches 99.9% of C_{eq} when the travel time equals the equilibrium time, and Da is Damköhler number defined as the ratio of the mean fluid travel time T_f (T) to the time required to reach chemical equilibrium T_{eq} (T):

$$Da = T_f / T_{\mathrm{eq}}. \tag{19}$$

The equilibrium time for chemical weathering T_{eq} is the ratio of equilibrium concentration C_{eq} to the reaction rate R_n (M·L^{-3}·T^{-1}) for the reaction of interest. The reaction rate can be estimated as the maximum reaction rate $R_{n,\max}$ (M·L^{-3}·T^{-1}) times the fraction of fresh minerals f_w, which is assumed to be 1 in this study. We use $C_{\mathrm{eq}} = 380\mu$mol/L and $R_{n,\max} = 1,085\mu$mol/L/year, representative of the dissolution of SiO_2. Thus, $T_{\mathrm{eq}} = C_{\mathrm{eq}}/(R_{n,\max} f_w) = 0.35$ year. The readers are referred to supporting information in Maher and Chamberlain (2014) for details about these parameters.

Note that we are aware of the fact that the parameter values for chemical weathering should be determined by the hydrogeologic conditions of the watershed and the properties of minerals of interest. But our modeling in this study is not aimed at reproducing the hydrogeochemical processes in any specific watershed. The chemical equilibrium time of the dissolution of SiO_2 is within the range of modeled baseflow age in our model configurations. So it can be used to illustrate the role of chemical equilibrium in influencing the effect of capturing/missing topographic features on modeled solute output from subsurface chemical weathering.

Table 1
Total Baseflow (m^3/s, Equation (10)) Generated Within the Watershed as a Function of Topographic Complexity Level (Rows) and Geological Heterogeneity Represented by the Extinction Depth d_e (Columns)

TCL / d_e	∞	1,000 m	333.3 m	100 m	33.3 m	10 m	3.3 m
1S	2,218	1,978	1,685	1,189	698	296	109
1SR	2,224	1,985	1,692	1,196	704	300	112
2S	2,223	1,985	1,692	1,198	706	302	112
2SR	2,239	2,000	1,707	1,211	719	311	117
3S	2,234	1,996	1,705	1,211	719	311	117
3SR	2,272	2,034	1,742	1,246	750	332	127
4S	2,253	2,016	1,723	1,229	734	321	122
4SR	2,298	2,060	1,768	1,273	774	349	136
$\Delta Q_{b,T}^{cl,\max}$ (%)	3.5	4.0	4.7	6.5	9.8	15.1	19.4

Note. $\Delta Q_{b,T}^{cl,\max}$. The maximum difference of total baseflow relative to TCL4SR for each d_e.

Subsurface chemical weathering from watersheds with too small or too large chemical equilibrium time, as compared to the baseflow age, will be controlled by only the chemical equilibrium time or the baseflow age, and will not be able to illustrate the interacting effect of baseflow age and chemical equilibrium time.

We assume that the residence time distribution for water discharging to the stream network from each nodal point in our model can be approximated as exponential with mean residence time $T_f = A_j$. Effectively, this approach treats the system as a collection of flow tubes, and it allows us to test the effect of TCL on modeled weathering at the watershed scale with a parsimonious and insightful modeling approach. Mean solute concentration for each stream reach is calculated as the flux-weighted average of solute concentrations of water draining from all nodal points into that stream:

$$C_{b,\omega,i}^{cl} = \frac{1}{Q_{b,\omega,i}^{cl}} \sum_{j \in \Omega_{node}^{\omega,i}} |Q_j| C_j. \tag{20}$$

Results for models with different TCLs are compared in the same manner as equations (12) and (13) to evaluate the effects of mesh TCL on stream chemistry.

3. Results and Discussion

3.1. The Effect of Capturing the Stream Network

In this subsection, we explain in detail the effect of capturing the stream network on the net baseflow generated in the simulation domain, groundwater flushing intensity, the flow rate, mean age, and solute concentration of baseflow into stream reaches of different Horton-Strahler stream orders. The effects are quantified using metrics described in section 2.2. The effect of capturing the stream network differs for streams of different orders and is influenced by the geological heterogeneity.

3.1.1. Net Flux Generated Within the Domain

For each geological heterogeneity scenario (i.e., distribution of hydraulic conductivity K), the net baseflow increases as the mesh is refined along more stream or ridge features (see Table 1, which summarizes the net baseflow generated within the simulation domain). The higher the level of refinement along the stream network and ridges, the higher the hydraulic head gradient driving flow, and the higher the groundwater flow rate and baseflow into streams. Similarly, mesh refinement allows the model to capture more local-scale flow paths, increasing the net baseflow and in particular the contribution of younger water (described in detail in sections 3.1.4 and 3.2.3).

Decreasing the extinction depth for hydraulic conductivity enhances the effect of model TCL on the net baseflow within the simulation domain (Table 1). The change of net baseflow for the whole domain due to mesh TCL is less than 10% for extinction depths $d_e \geq 33.3$ m (Table 1). This indicates that for watersheds with fractured and permeable bedrock, which tend to have relatively deep extinction depths, the effect of mesh TCL on net groundwater discharge is likely to be relatively small. On the other hand, for watersheds where the permeability of the bedrock decreases relatively fast with depth ($d_e \leq 10$ m), the effect of mesh TCL on net groundwater discharge within the domain can be significant (of the order of 20%, Table 1).

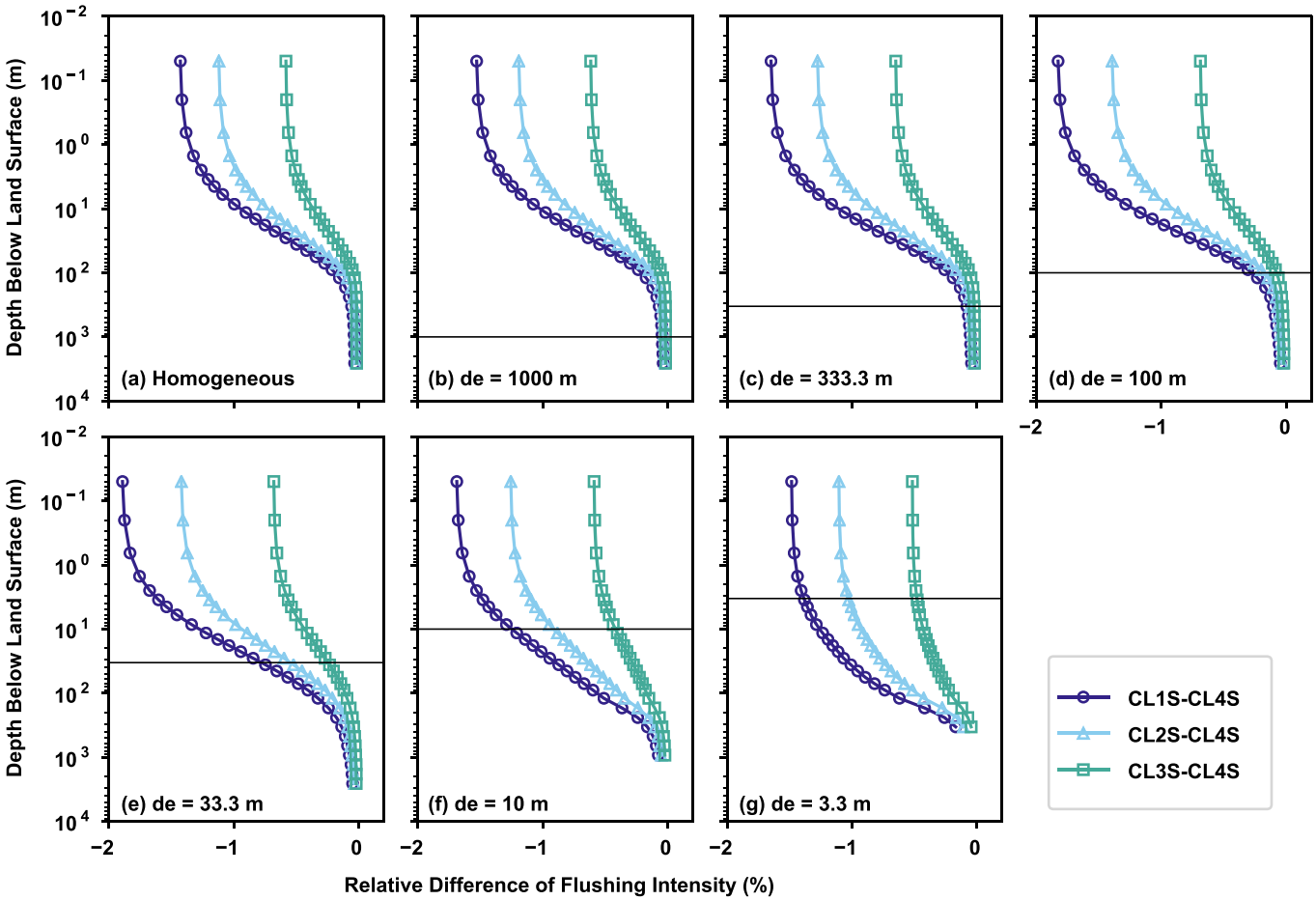

Figure 3. Effect of mesh refinement along the stream network (line colors) on the flushing intensity for different geological heterogeneity scenarios (a–g). Flushing intensity of groundwater flow from models with meshes refined along a different amount of streams (TCL1S, TCL2S, and TCL3S) is compared to the result from the model with mesh refined along all the stream features (TCL4S). Horizontal axis corresponds to the relative difference of flushing intensity with the highest level of refinement (TCL4S) as a reference. A negative relative difference implies that missing lower-order streams in the mesh decreases the capacity of the flow system to transport water mass by advection at different depths. The horizontal solid lines represent the extinction depths.

The maximum relative difference ($\Delta Q_{b,T}^{cl,\max}$ in Table 1) of net baseflow caused by mesh TCL for each geological heterogeneity scenario is relatively small and, in particular, is of the order of magnitude of uncertainty in typical estimates of recharge (e.g., Healy, 2010; Scanlon et al., 2002) or baseflow separation (e.g., Santhi et al., 2008) from observations or modeling. From a practical perspective, for the same geological heterogeneity, all the TCLs produce approximately the same total amount of baseflow; that is, the net response of the model is indistinguishable within the range of likely uncertainty. In the following sections we describe in detail the impact of mesh TCL on the internal characteristics of the flow and transport system.

3.1.2. Flushing Intensity of Groundwater Flow

The magnitude of flushing intensity, a metric of the net capacity of the groundwater flow system to move water, decreases with depth (Figure S1 in the supporting information). And the flushing intensity is higher for models with deeper extinction depths (Figure S1). The lack of mesh refinement along lower-order streams (e.g., TCL1S, TCL2S, and TCL3S) results in a negative bias of flushing intensity when compared with the model refined along all streams (TCL4S), and this is consistent for all geological heterogeneity scenarios (Figure 3). The bias is highest at the surface and progressively decreases with depth. Here it is important to note that the bias is always smaller than 2%.

Geologic heterogeneity influences the effect of TCL on the flushing intensity. For extinction depths deeper than 10 m (Figures 3a–3e), decreasing extinction depth enhances the bias of flushing intensity caused by mesh TCL, and most of the enhancement occurs above the extinction depth (horizontal lines in Figure 3). This is because shallow extinction depth suppresses the development of the deep intermediate and regional

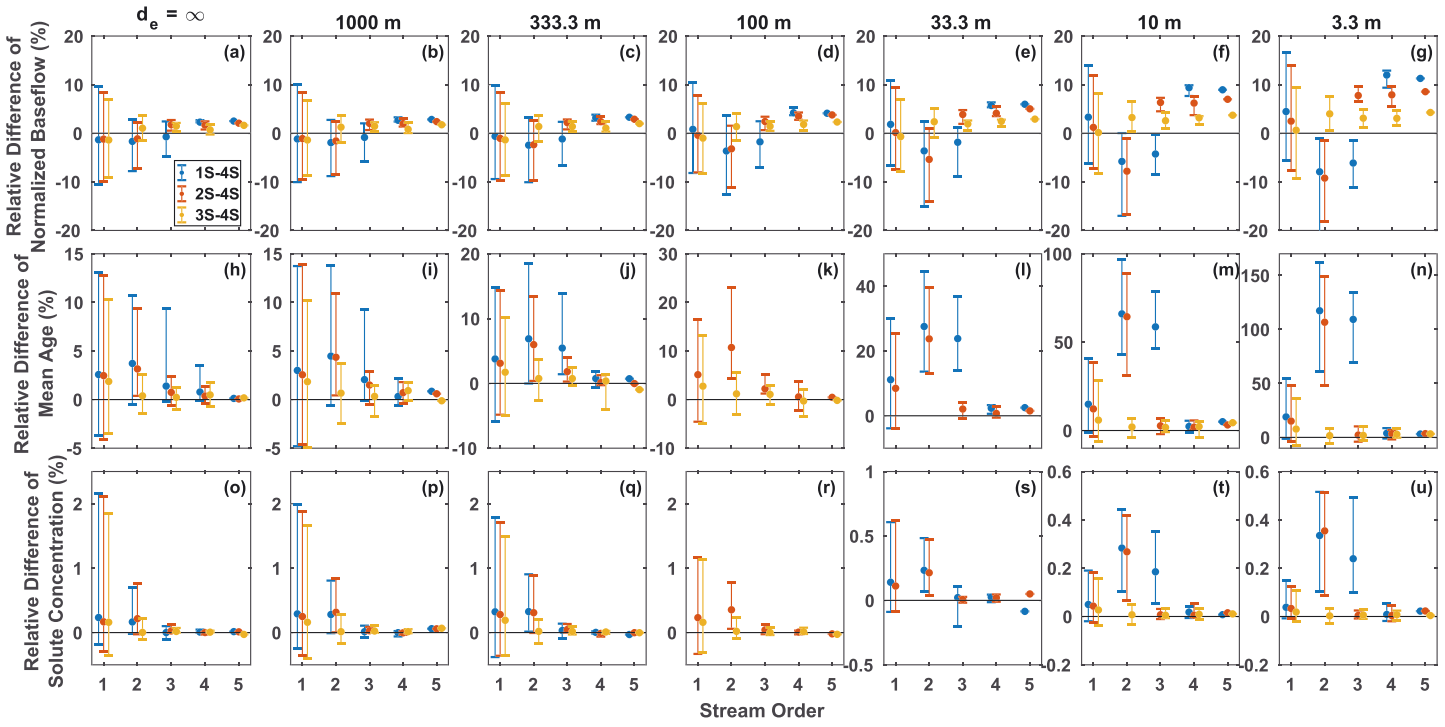

Figure 4. Effect of capturing the stream network on the normalized flow rate, mean age, and solute concentration (SiO_2 from subsurface chemical weathering, as an example) of baseflow draining into streams of different orders. The relative difference of each metric is calculated as described in section 2.2. Results from models with meshes capturing a different amount of stream features (TCL1S, TCL2S, and TCL3S) are compared to the result from the model with mesh capturing all the stream features (TCL4S). A positive relative difference means more baseflow (a–g), larger baseflow age (h–n), or higher baseflow solute concentration (o–u) for a stream in a model with mesh TCL1S, TCL2S, or TCL3S compared to the same stream in the model with mesh TCL4S. The median value (dots) and interquartile range (error bars) of relative differences for streams of each order are plotted. Columns show results for simulations with different extinction depths, as indicated by the label at the top of the figure. (Mean age simulations using model with *de* = 100 m, TCL1S and model with *de* = 33.3 m, and TCL3S have convergence issues and are not presented.)

groundwater flow paths and enhances groundwater circulation along the shallow local flow paths (Cardenas & Jiang, 2010). The shallow, local flow paths are affected more by capturing local water table undulations than the deep regional flow paths. However, as the extinction depth further decreases (Figures 3f and 3g), the bias of flushing intensity above the extinction depth becomes smaller, but the bias below the extinction depth becomes larger than for the deeper extinction depth cases.

3.1.3. Groundwater Circulation and Baseflow Generation

In addition to the effect of mesh TCL on the integrated metrics, including the net flux of baseflow generated within the simulation domain and flushing intensity of groundwater flow through layers, significant and different spatially distributed effects can be found for baseflow, its mean age, and its solute concentration for streams of different orders (Figures 4, 5, S7, and S8), further reflecting the effect of mesh TCL on the nesting and spatial distribution of the groundwater flow paths. The flow rate, mean age, and solute concentration of baseflow into each stream from models with TCL*icl*S (*icl* = 1, 2, 3) are compared to those from the model with TCL4S, following the approach described in section 2.2. We discuss in detail the effect of capturing the stream network on baseflow (Figures 4a–4g and 5) in this subsection, and on baseflow mean age (Figures 4h–4n and S7) and solute concentration (Figures 4o–4u and S8) in the next two subsections.

Generally, for models with meshes that capture less stream features, baseflow into higher-order streams (ω = 3, 4, 5) are dominated by positive bias (more baseflow), while baseflow into lower-order streams (ω = 1, 2) are dominated by negative bias (less baseflow) (Figures 4a–4g and 5). This can be explained by the ability of high TCL meshes to capture topographic variability at the local scale and therefore capture high-frequency spatial variations in water table that ultimately drives water through local-scale flow paths and generate more baseflow in low-order streams (Zijl, 1999). Meanwhile, flow through intermediate and regional flow paths is reduced, resulting in less baseflow into the higher-order streams.

In particular, for the fifth-order stream, baseflow is biased higher in models with TCL*icl*S (*icl* = 1, 2, 3) as compared to model with TCL4S (Figures 4a–4g and 5). As more streams of order lower than 4 are captured

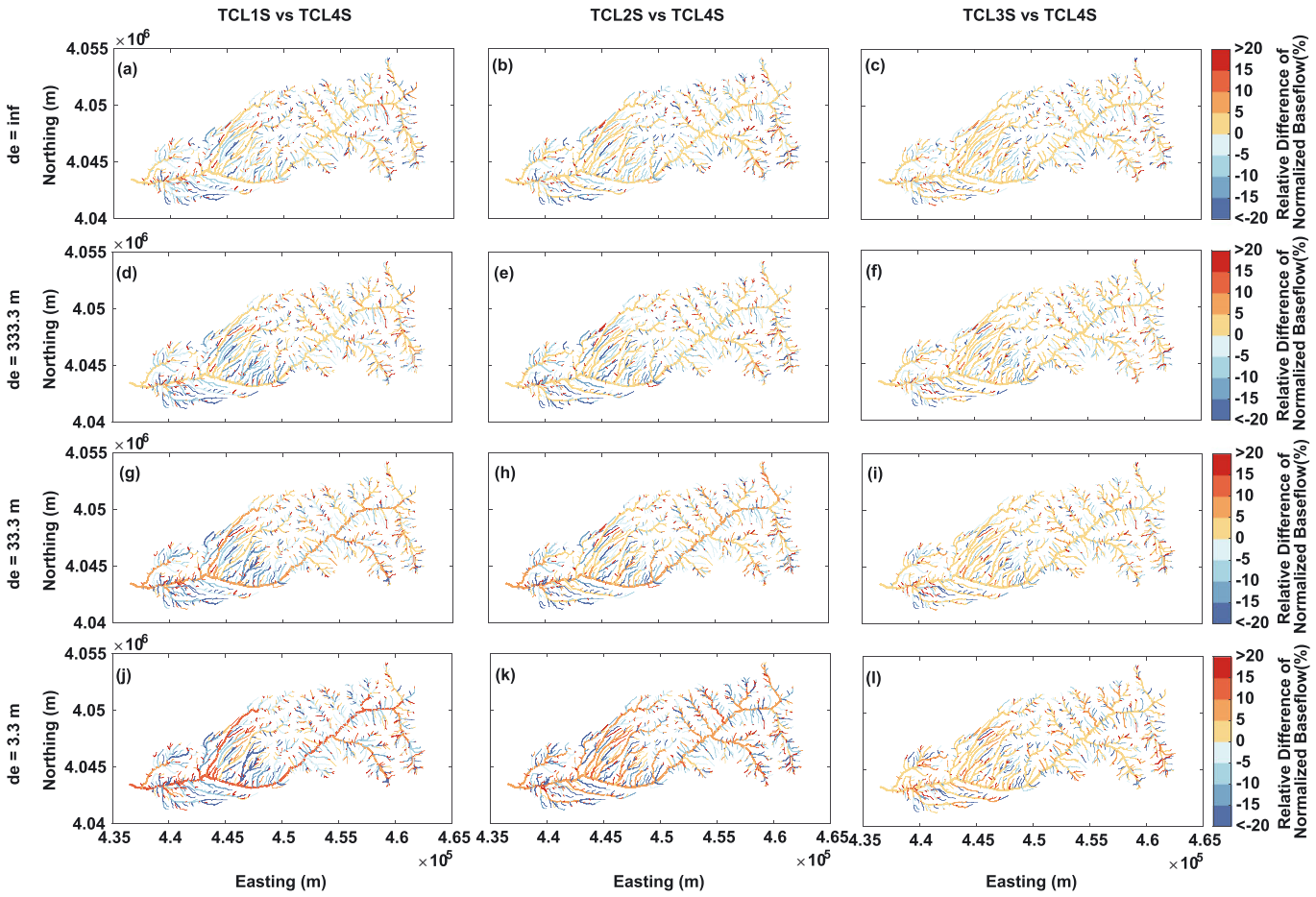

Figure 5. Effect of capturing the stream network on the normalized baseflow draining into streams. The relative difference is calculated as described in section 2.2. Columns show results from models with meshes capturing a different amount of stream features (TCL1S, TCL2S, and TCL3S) compared to the result from the model with mesh capturing all the stream features (TCL4S), as indicated by the label at the top of the figure. Rows show results for simulations with different extinction depths, as indicated by the label at the left of the figure.

by the model (from TCL1S to TCL4S), the positive bias of baseflow into the fifth-order stream decreases (Figures 4a–4g), which indicates that capturing lower-order streams besides the main stream can have a significant effect on simulated baseflow into the main high-order stream. A similar effect can be observed for fourth-order streams. For third-order streams, when they are not captured by the model (TCL1S), baseflow into these streams is biased lower (Figures 4a–4g). Once they are captured by the model (TCL2S and TCL3S), baseflow into these streams is biased higher, and the interquartile range (IQR) of the bias is smaller. The median bias decreases from the model with TCL2S to the model with TCL3S, as TCL3S better captures second-order streams than TCL2S. Following the same thread of thought, baseflow into second-order streams is biased lower when second-order streams are not captured by the model (TCL1S and TCL2S) and is biased higher with smaller bias IQR when second-order streams are captured by the model (TCL3S).

These biases due to mesh resolution along the stream network are consistent for all geologic heterogeneity scenarios, but their magnitude increases with shallower extinction depths (Figures 4a–4g and 5). These detailed but nuanced effects can be summarized. First, if streams of order ω are not captured in a model with TCL*icl*S, then baseflow into streams of order ω from the model with TCL*icl*S is less than baseflow into streams of order ω from models with meshes that capture streams of order ω. The more streams of order other than ω are captured in model with TCL*icl*S, the less baseflow flows into streams of order ω. Second, if streams of order ω are captured in the model, then the less streams of order other than ω are captured simultaneously, the more baseflow flows into streams of order ω. Third, refining the mesh along streams make those streams more "competitive" in draining groundwater to generate baseflow. Fourth, refining the mesh along streams

of certain order affects not only the baseflow into streams of that order but also the subsurface flow paths across all scales and baseflow into streams of other orders.

For first-order streams, the magnitude of median bias is relatively small, especially for models using hydraulic conductivities with extinction depths $d_e \geq 100$ m (Figures 4a–4d). The more streams of stream order higher than one are captured by the mesh, the less the groundwater discharge into first-order streams (e.g., Figures 4c, from TCL1S to TCL3S). This is consistent with the effect of mesh TCL on baseflow into second- to fifth-order streams, as described above. However, the predominant bias of baseflow into first-order streams changes from negative to positive, with baseflow into second- and third-order streams biased more negatively, as the extinction depth of hydraulic conductivity decreases (Figures 4a–4g). In models with relatively deep extinction depth (Figures 4a–4c), baseflow into first-order streams is predominantly biased lower when first-order streams are not captured by the mesh. This is consistent with our explanation that ignoring local-scale topographic variability along these first-order streams smooths the high-frequency fluctuations in water table, driving less water through local-scale, short flow paths that drain into first-order streams. For extinction depths ≤ 10 m (Figures 4f and4g), ignoring first-order streams (TCL1S, TCL2S, and TCL3S) results in positive bias of baseflow into these streams, as compared to the model with a mesh capturing first-order streams (TCL4S) .

We further created 2-D cross-sectional models to illustrate the effect of mesh refinement along streams on groundwater flow paths and baseflow into streams of different orders (Figure 6). The effect of mesh refinement on the baseflow into each stream from these 2-D conceptual models, as shown in Figure 6a, is consistent with the results analyzed above for the 3-D model. Figure 6e shows the flow streamlines for the model with mesh refined at all stream locations, as marked by the solid triangles, and is used as a reference case for comparison. When the first-order stream is not refined in the mesh, groundwater draining into the first-order stream through local flow paths decreases, as shown by the sparser streamlines toward the first-order stream (Figure 6d). When the mesh further misses the second-order stream, groundwater draining into the second-order stream decreases (Figures 6a and 6c). And similar effect can be observed for the third-order streams (Figures 6a and 6b). We can also find that when the mesh is not refined at lower-order streams, streamlines penetrate deeper and become longer (e.g., comparing Figures 6b and 6e), and this causes the baseflow age to be biased older (see section 3.1.4 for more details).

Returning to the 3-D model, the IQR of the bias in baseflow is also affected by the model TCL. Low IQR means that the biases are more consistent and closer to the median bias. Generally, the IQR of the bias decreases as more topographic features are captured by the mesh (Figures 4a–4g). As expected, the bias IQR for streams of order ω decreases significantly once the streams of order ω are captured by the mesh. Changes of bias IQR indicate that different streams of the same order have different sensitivity to model TCL in terms of baseflow generation. The bias IQR increases with decreasing stream order. This is because of the higher number of lower-order stream reaches and that lower-order streams tend to gain water from shallower flow paths, which are influenced more by capturing topographic variability (Hale & McDonnell, 2016; McGuire et al., 2005).

Decreasing extinction depth of hydraulic conductivity enhances the biasing effect of low TCL. For example, systems with shallow extinction depth are characterized by higher values of median bias, larger differences of the bias among model TCLs, and larger bias IQR (Figures 4a–4g). This is consistent with previous modeling effort where decreasing of the extinction depth enhances the shallow circulation by forcing more water into shallow and local flow paths (Cardenas & Jiang, 2010).

Although we use an idealized model where groundwater table is assumed to be a replica of land surface elevation (Haitjema & Mitchell-Bruker, 2005), our findings are consistent with previous studies. For example, Vivoni et al. (2005) find that a coarse model smooths the terrain, reduces the gradients in the water table topography, decreases groundwater exfiltration and baseflow discharge, increases surface runoff, and leads to a shift from slow subsurface runoff to quick surface flow. Although the model used by Vivoni et al. (2005) involves surface flow, vadose zone, and shallow groundwater flow processes, their modeling does not attempt to reproduce deep groundwater flow processes and the effect of capturing stream and ridge features on baseflow discharge at different scales. For some watersheds, deep groundwater flow has been found to be important for baseflow generation and stream chemistry evolution and is responsible for producing the simple emergent scaling relationship between groundwater contribution to stream flow and drainage area (discussed in detail in section 3.3; e.g., Frisbee et al., 2011, 2017; Peralta-Tapia et al., 2015). Thus, understanding how watershed topographic complexity affects groundwater contribution to streamflow across scales is important to

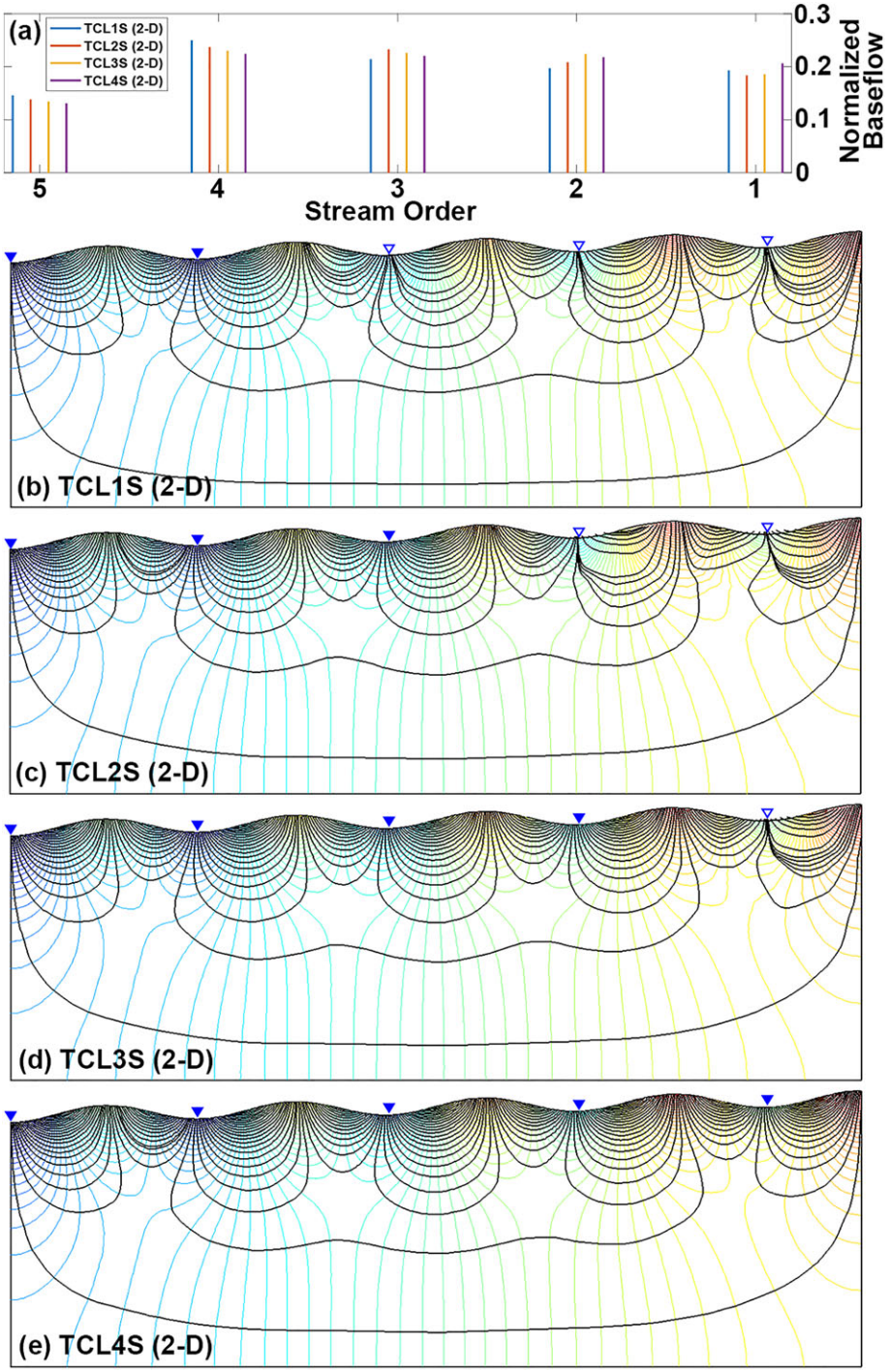

Figure 6. The 2-D cross-sectional Toth model illustrating the effect of capturing the stream network on (a) baseflow into streams of different orders, caused by (b)–(e) groundwater reallocation through the nested flow paths of different scales. (a) The proportion of total baseflow into each catchment. (b)–(e) The streamlines (black solid) for each 2-D model. The flow between each pair of adjacent streamlines is the same throughout the domain, giving denser streamlines where the magnitude of the flow field is high. Blue triangles in (b)–(e) mark the locations of streams, with solid (hollow) triangles marking the streams where mesh is (not) refined. (e) Mesh scenario TCL4S where streams of all orders are used for mesh refinement. (b) to (d) Mesh scenarios with reduced topographic complexity levels (TCLs). This conceptual model corresponds to systems with relatively deep extinction depths (e.g., $d_e = 1,000$ m).

building parsimonious watershed scale hydrological models (McDonnell et al., 2007) and improving hydrological predictions in data-sparse regions (Sivapalan et al., 2003).

Furthermore, the biasing effect in baseflow due to the lack of refinement along low-order streams is expected to be particularly important when the model is used to reproduce ecological processes and evapotranspiration, given the role that the spatial distribution of stream flow generation can have on riparian vegetation. A model with low TCL generates more water discharging into the main stream and could make more plants concentrate along main stream riparian zone and fewer plants along tributaries, thus influencing the spatial distribution of evapotranspiration and further biasing the simulated stream flow spatial distribution and water balance.

3.1.4. Baseflow Mean Age

Generally, the mean age of baseflow for models with low TCL meshes (i.e., capturing fewer lower-order streams) is predominantly biased older for all stream orders (Figures 4h–4n; note that the vertical age scale varies with d_e; Figure S7). However, capturing streams of order ω can significantly reduce the bias for those streams (e.g., from TCL3S to TCL4S for stream order of 1, from TCL2S to TCL3S for stream order of 2, in Figure 4h). This is consistent with the effect of mesh TCL on the normalized flow rate. Capturing high-frequency topographic features through mesh refinement drives more water through shorter and shallower local-scale subsurface flow paths and less water through longer and deeper regional-scale flow paths. This leads to more young water flowing into lower-order streams and less old water flowing into higher-order streams, which means that baseflow mean age for models with high TCL is dominantly younger than that for models with low TCL.

For the first-order streams, age of baseflow for models with TCL1S, TCL2S, and TCL3S is predominantly biased older than that for model with TCL4S (Figures 4h–4n). Comparing TCL1S, TCL2S, and TCL3S to TCL4S, we find that the most significant bias is caused by not refining the mesh along first-order streams when second- to fifth-order streams are captured in the mesh (from TCL4S to TCL3S). This means that when higher-order streams are captured in the mesh, the marginal effect of missing first-order streams on the simulated baseflow mean age can be significant for first-order streams. This has important implications for modelers because typical modeling efforts ignore the low-order streams to reduce computational burden at the expense of significant biases in mean age for headwater streams. Missing streams of order higher than 1 (TCL2S and TCL1S) gradually leads to older and older baseflow, but the increase of bias is not as significant as that caused by missing first-order streams (from TCL4S to TCL3S). The IQR of the bias also increases from TCL3S to TCL1S, indicating that age of baseflow into first-order streams can also be affected by not refining the mesh along higher-order streams, which leads to more extreme bias in some stream reaches. This can be informative to interpret modeling results since it tells us that some stream reaches may have a much higher or lower bias than the median, suggesting the need of careful interpretation depending on the modeling purpose. These patterns in bias are consistent for geologic heterogeneities, but the magnitude of both the median value and the bias IQR increase as the extinction depth decreases (recall that the vertical age scale varies with d_e).

For second-order streams, the effect of mesh TCL on baseflow age is similar to that for first-order streams (Figures 4h–4n). In brief, capturing more lower-order streams from TCL1S to TCL3S decreases the median value and the IQR of the bias. The most significant decrease in the median bias value and its IQR occurs when we refine the mesh along second-order streams (from TCL2S to TCL3S). Missing first-order streams does not cause as much bias of mean age for second-order streams (from TCL4S to TCL3S). The bias effect is similar for different geologic heterogeneity scenarios but is enhanced for hydraulic conductivity with shallow extinction depth. Similar results can be found for third- and fourth-order streams.

For the fifth-order stream, the bias of baseflow age caused by mesh TCL is within 5% for all geologic heterogeneity scenarios but increases as the extinction depth of hydraulic conductivity decreases (Figures 4h–4n and S3). For the model using homogeneous hydraulic conductivity, there is little bias caused by mesh TCL. As the extinction depth of hydraulic conductivity decreases, but still deeper than 3.3 m, missing more and more lower-order streams (TCL3S, TCL2S, and TCL1S) biases the baseflow older (Figures 4h–4m). When the extinction depth of hydraulic conductivity is very shallow ($d_e = 3.3$ m), missing streams higher than order of 1 has no significant impact on the age (TCL1S, TCL2S, and TCL3S), and the bias is mainly caused by missing the first-order streams (from TCL4S to TCL3S, Figures 4n and S3n). The groundwater circulation system in a model with shallow extinction depth for hydraulic conductivity is dominated by shallow and local flow paths (Cardenas & Jiang, 2010) and suppresses the intermediate and regional flow paths. This kind of groundwater system is more sensitive to capturing first-order streams than to capturing higher-order streams. These

observations indicate that for modeling in watersheds where bedrock permeability decreases fast with depth, baseflow into the main stream (fifth order in this study) might not be sensitive to mesh refinement along streams higher than order of 1 but unexpectedly could be sensitive to mesh refinement along streams of order of 1. The common modeling practice in which the mesh is refined along only high-order streams can lead to significant bias of the age of baseflow into the main stream. So the modeler should be more cautious to the effect of mesh resolution along headwater streams for watersheds dominated by shallow flow.

The effect of model TCL on baseflow mean age has similar patterns for hydraulic conductivity fields with different extinction depths. Capturing detailed local topographic variability affects baseflow age more for models with shallower extinction depths, as evidenced by the increased magnitude of the median bias and the bias IQR (Figures 4h–4n and S4h–S4n). In this case, systems with shallow extinction depth force groundwater to dominantly flow through the upper part of the modeling domain and the effects of capturing local topographic variability on the spatial distribution of groundwater discharge are enhanced. This is also consistent with field observations (e.g., Hale & McDonnell, 2016). Watersheds with shallow and relatively impermeable bedrock are dominated by shallow flow paths. Streamflow mean transit time in these watersheds is more correlated with topographic indices, including median flow path gradient and median path length, than watersheds with fractured permeable bedrock (Hale & McDonnell, 2016). If we further look at the same plots with a fixed scale for y axis (Figures S3 and S4), we can find that the influence of geologic heterogeneity is more pronounced for extinction depths $\leq$ 33.3 m, and the age of baseflow into streams of order of 1 to 3 are the most sensitive to mesh TCL (Figures S3l–S3n and S4l–S4n). This indicates that the age of baseflow in lower-order streams in watersheds with tight bedrock, which have shallow extinction depth, is the most sensitive to capturing stream features in the mesh.

3.1.5. Solute Concentration of Baseflow

The concentration of solute (SiO_2 as an example in this study) from subsurface chemical weathering is a function of Damköhler number (equation (19)), which is proportional to groundwater residence time. As groundwater residence time approaches the chemical equilibrium time for the reaction of interest, the solute concentration approaches the equilibrium concentration (equation (18)). The effect of mesh TCL on baseflow solute concentration is consistent with that on baseflow mean age; that is, solute concentration is biased higher when mean age is biased older due to the lack of mesh refinement along low-order streams (Figures 4o–4u and S8). On the other hand, for the high-order streams, which get very old baseflow, the bias effect of mesh TCL is diminished due to the constraint of chemical equilibrium concentration (equation (18)).

For first-order streams, low TCL results in longer groundwater residence time (i.e., time available for reactions to occur) and thus higher solute concentration. The effect of mesh TCL is most significant when the mesh is not refined along first-order streams (from TCL4S to TCL3S). A similar effect can be found for second-order streams. But the median bias is within 1%, and the IQR of the bias is within 2%, which means that the simulated baseflow solute concentration is less sensitive to mesh TCL than the flow rate and mean age. This is due to two reasons. First, for all the model scenarios, more than 60% of stream reaches have baseflow mean age larger than T_{eq}, which is 0.35 year from the parameters used in this study. The difference in mean age larger than T_{eq} cannot be reflected in solute concentration due to the limit of chemical equilibrium. Second, solute concentration is not sensitive to groundwater residence time. In fact, a relative large change in groundwater residence time results in a smaller change in solute concentration, especially when groundwater residence time is closer to T_{eq} (supporting information Text S1 and Figure S2).

Mesh TCL has a stronger effect on solute concentration in low-order streams ($\omega = 1, 2$) for models with deep extinction depth (e.g., the homogeneous K). This is evidenced by the magnitude of the median value and the IQR of the bias (Figures 4o–4u). Simulated solute concentration is less affected by mesh TCL in models with shallow extinction depth (Figures 4s–4u and S8) because baseflow mean age generated by these models is relatively old (Figures 9m and 9n and 9t and 9u) and the chemical weathering reaches equilibrium well before the groundwater discharges into the streams. In fact, for $d_e \leq$ 33.3 m, more than 99% stream reaches have baseflow mean age larger than T_{eq}. The high-order streams ($\omega = 3, 4, 5$) also have an unbiased median concentration when the model TCL changes, because higher-order streams gain water from longer regional flow paths with relatively long residence time during which the chemical weathering reaches equilibrium before the groundwater discharges into streams.

3.2. The Effect of Refining the Mesh Along Ridges

In this subsection, we focus on the effect of capturing the ridges of different order catchments. Even if the mesh is refined along streams of a given order, the lack of refinement along the ridges defining stream

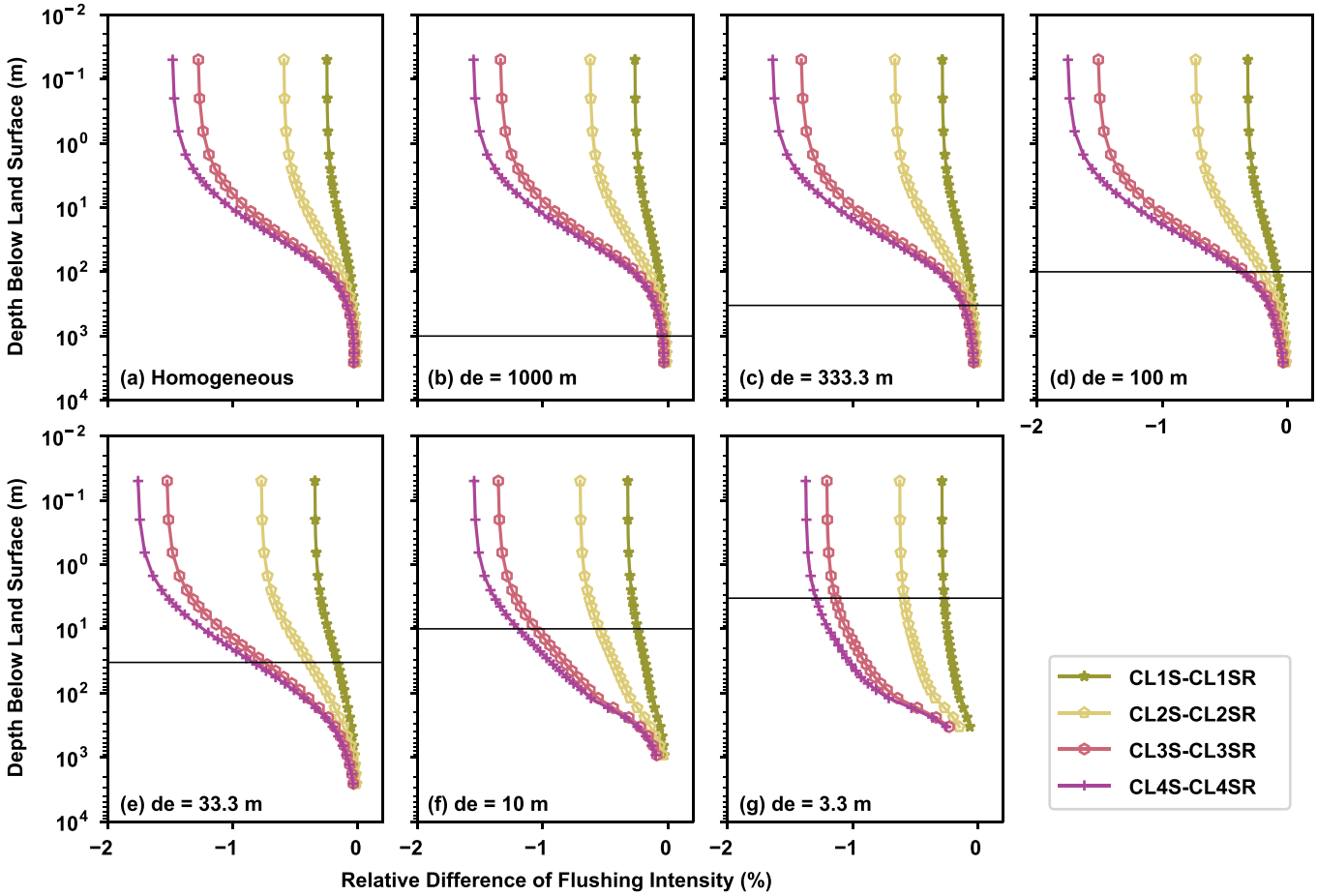

Figure 7. (a–g) The effect of having the mesh capture the ridges in addition to capturing streams on the flushing intensity of simulated groundwater flow and the influence of geological heterogeneity. Flushing intensity of groundwater flow from models with meshes capturing streams only (TCL1S, TCL2S, TCL3S, and TCL4S) is compared to that from models with meshes capturing streams as well as ridges of the corresponding order (TCL1SR, TCL2SR, TCL3SR, and TCL4SR). A negative relative difference implies that missing ridges corresponding to streams of the same orders in the mesh decreases the capacity of the flow system to transport water mass by advection at different depths. Different subplots show results from simulations using hydraulic conductivity fields with different extinction depths, as marked by the solid horizontal lines.

drainage area can cause important biases. Similar to previous sections, we focus on one integrated metric, flushing intensity, and three spatially distributed metrics, baseflow, baseflow mean age, and the concentration of weathering products in baseflow.

3.2.1. Groundwater Flushing Intensity

If the mesh is refined along streams of certain orders, but not along the corresponding ridges, then an additional negative bias of flushing intensity can occur (Figure 7). Recall that in this section, the bias is calculated between the model with mesh capturing streams only and the model with mesh capturing streams as well as ridges of the corresponding order (equation (16)). The bias is largest at the landscape surface and decreases with depth, but its magnitude is less than 2% (Figure 7). The pattern of the bias is consistent under different geological heterogeneity conditions, but the magnitude of the bias above the extinction depth is enhanced when the extinction depth decreases but deeper than or equal to 33.3 m (Figures 7a–7e). When the extinction depth further decreases (shallower than or equal to 10 m), the magnitude of the bias decreases above the extinction depth and increases below it (Figures 7f and 7g). If only high-order streams are used for mesh refinement, then the bias caused by missing the corresponding high-order ridges in the mesh is relatively small (TCL1S vs. TCL1SR). However, when streams of all orders are used for mesh refinement, the bias caused by not refining the mesh along corresponding ridges is the largest (TCL4S vs. TCL4SR).

3.2.2. Groundwater Circulation and Baseflow Generation

Not refining the mesh along the ridges corresponding to the order of streams used for mesh refinement leads to negative bias of baseflow for first-order streams, and positive bias for streams higher than order of 1

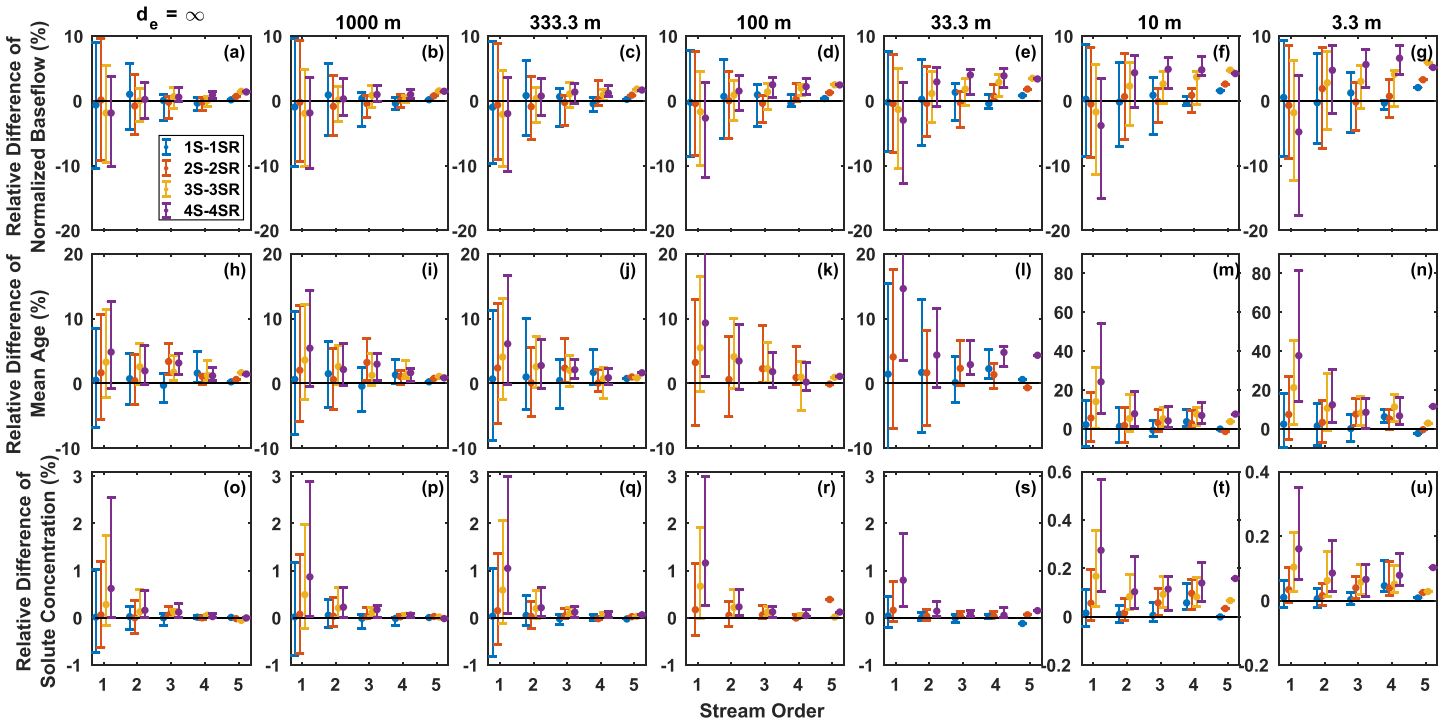

Figure 8. Effect of capturing the ridges in meshes in addition to the stream network on the normalized flow rate, mean age, and solute concentration (SiO_2 from subsurface chemical weathering, as an example) of baseflow draining into streams of different orders. Results from models with meshes capturing only stream features (TCL1S, TCL2S, TCL3S, and TCL4S) are compared to the results from models with meshes capturing streams as well as ridges of the corresponding order (TCL1SR, TCL2SR, TCL3SR, and TCL4SR). A positive relative difference means more baseflow (a–g), larger baseflow age (h–n), or higher baseflow solute concentration (o–u) for a stream in a model with mesh TCL1S, TCL2S, TCL3S, or TCL4S compared to the same stream in the model with mesh TCL1SR, TCL2SR, TCL3SR, or TCL4SR, respectively. Columns show results for simulations with different extinction depths, as indicated by the label at the top of the figure. (Mean age simulation using model with *de* = 100 m, TCL1S, and model with *de* = 33.3 m, TCL3S, have convergence issues and are not presented.)

(Figures 8a–8g and S9). Decreasing extinction depth increases both the magnitude of the median bias, and the bias IQR, which is consistent with the effect of mesh TCL on the flushing intensity.

For first-order streams, the biasing effect is more significant when first- and second-order ridges are not captured by the mesh, but, respectively, first- and second-order streams are captured (TCL3S vs. TCL3SR, and TCL4S vs. TCL4SR). For second-order streams, when the extinction depth $d_e \geq 333.3$ m, the bias is the most significant when fourth- and fifth-order ridges are not used for mesh refinement (TCL1S vs. TCL1SR). But when the extinction depth $d_e \leq 100$ m, the bias is more significant when lower-order ridges are not used for mesh refinement (e.g., TCL4S vs. TCL4SR). For third-order streams, bias is minimal between the model TCL2S and TCL2SR, which means that when streams of order 5, 4, and 3 are used for mesh refinement, the bias caused by not refining the mesh along corresponding ridges is small. For fourth- and fifth-order streams, the bias increases as the mesh captures more streams but no corresponding ridges (compare TCL1S to TCL4S with TCL1SR to TCL4SR).

Generally, ridges become important when corresponding streams are used for mesh refinement. The bias caused by not refining the mesh along the ridges is relatively small when only high-order streams are used for mesh refinement (e.g., TCL1S vs. TCL1SR). However, if lower-ordered streams are needed for mesh refinement in order to reduce the bias of simulated hydrological variables (section 3.1), then the mesh should also be refined along lower-order ridges to reduce the bias. Otherwise, the mesh capturing all the streams (TCL4S), which is used as a benchmark to evaluate the effect of missing streams, leads to large bias compared to the model with mesh capturing both the streams and ridges (TCL4SR). Although it might not be possible to refine the mesh along all streams and ridges due to the limit of computational capacity, the bias effect of missing these topographic features should be taken into account when interpreting the modeling results.

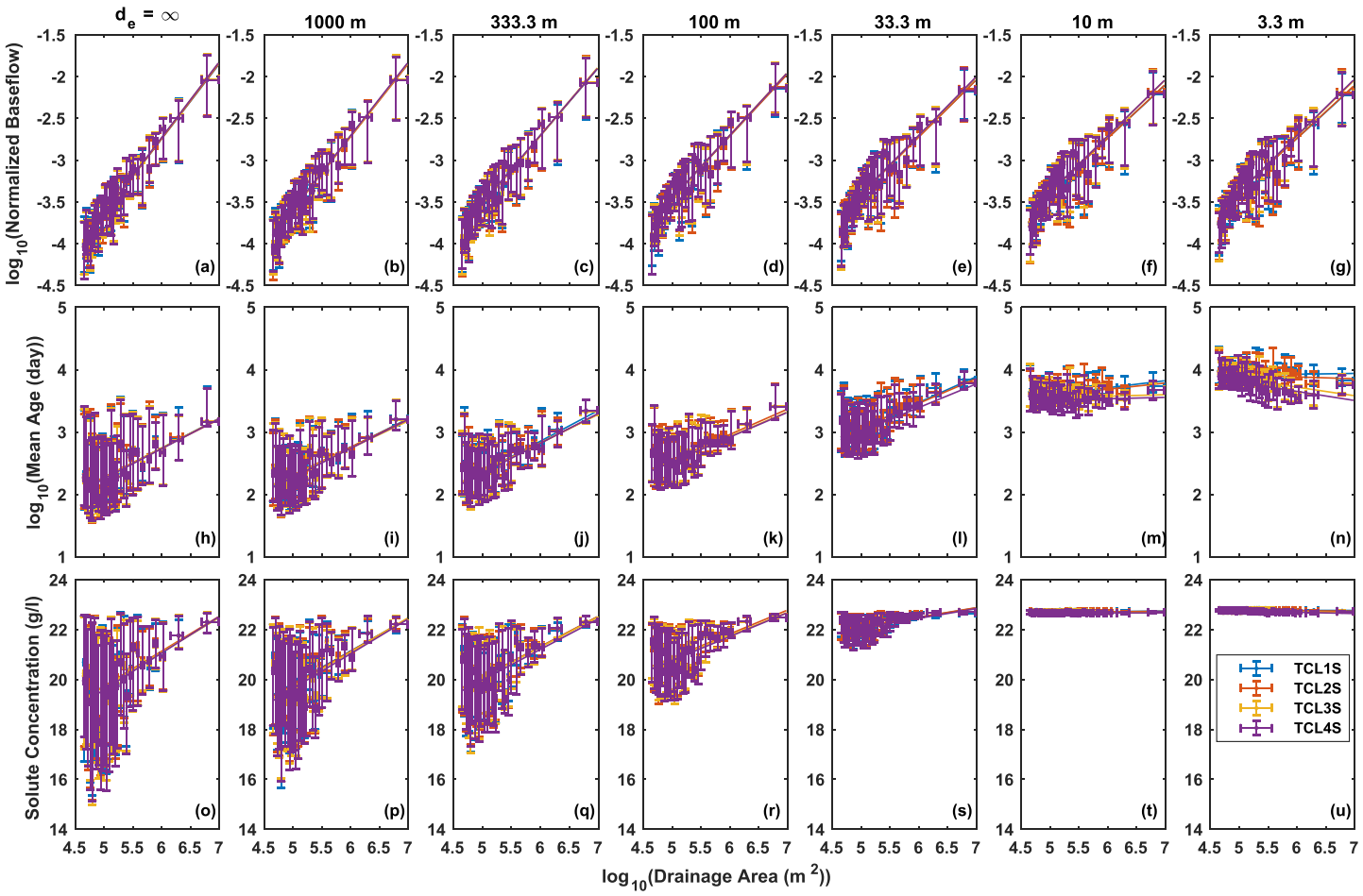

Figure 9. Scaling of normalized flow rate (a–g), mean age (h–n), and solute concentration (o–u) of baseflow with drainage area from models with meshes capturing streams only. The median value (dots) and interquartile range (error bars) of the variables within bins are plotted. Columns show results from simulations with different extinction depths, as indicated by the label at the top of the figure. (Mean age simulation using model with *de* = 100 m, TCL1S, and model with *de* = 33.3 m, TCL3S, have convergence issues and are not presented.)

3.2.3. Baseflow Mean Age

Capturing only streams but no ridges in the mesh biases the baseflow older (Figures 8h–8n and S10). This is consistent with the biasing effect on baseflow, because ignoring low-order ridges smooths the local-scale water table variability and causes more water to flow through longer, regional flow paths and less water to flow through shorter, local flow paths.

For first-order streams, baseflow is biased older and older when more and more streams are captured by the mesh but no corresponding ridges are captured (Figures 8h–8n, from TCL1S to TCL4S). For second-order streams, the most significant bias occurs when second-order streams are captured in the mesh but no ridges are captured (TCL3S vs. TCL3SR). A similar effect can be found for third-order streams. That is, when all the third-order streams but no third-order ridges are captured by the mesh, the bias is the most significant. For fifth-order streams, the bias gets larger when more ridges are not captured by the mesh. The effect of mesh TCL on baseflow mean age is consistent under different geologic heterogeneity conditions, but the effect is enhanced in terms of the median bias and the bias IQR as the extinction depth decreases.

3.2.4. Baseflow Solute Concentration

The effect of missing ridges in the mesh on baseflow solute concentration is consistent with that on baseflow mean age, since baseflow solution concentration is directly a function of baseflow mean age (equations (18) and (19)). The effect is more significant for lower-order streams and for models with shallow extinction depths, but the overall bias is within 3% (Figures 8o–8u and S11). Baseflow solute concentration is not as sensitive to capturing ridges in the mesh as are the flow rate and mean age, especially for high-order streams and for watersheds with rapidly decaying hydraulic conductivity and dominated by old baseflow.

3.3. The Effect of Mesh TCL on Baseflow Spatial Scaling

Some previous studies show that streamflow solute concentration increases with increasing accumulated drainage area, which is well explained by a 3-D catchment-mixing streamflow generation conceptual model with a topographically driven nested groundwater flow system (e.g., Frisbee et al., 2011; Peralta-Tapia et al., 2015). Other studies show that the plot of streamflow solute concentration versus drainage area is scattered with no significant trending at small scales but asymptotically approaches some median value as drainage area increases, which is explained by a shallow flow system and 2-D network-mixing streamflow generation conceptual model (e.g., Temnerud & Bishop, 2005; Uchida et al., 2005). Different forms of the scaling relationship between streamflow solute concentration and drainage area reveal different roles of groundwater in runoff generation under different topographic and geological conditions, with implications for modeling streamflow generation and water quality. Our synthetic models with relatively deep extinction depths (e.g., $d_e \geq 33.3$ m) correspond to the 3-D catchment-mixing streamflow generation conceptual model (Frisbee et al., 2011), and models with very shallow extinction depths (e.g., $d_e \leq 10$ m) correspond to the subsurface part of the 2-D network-mixing streamflow generation conceptual model (Frisbee et al., 2011; Temnerud & Bishop, 2005; Uchida et al., 2005). So we further test if our model can reproduce the spatial scaling relationships between baseflow, mean age, solute concentration and drainage area observed in natural watersheds, and the impact of mesh TCL on these scaling relationships. Understanding the influence of capturing topographic variability under different geological heterogeneity conditions on the large-scale emergent behavior of baseflow generation is a critical step toward building parsimonious watershed hydrological models (McDonnell et al., 2007; Sivapalan, 2003).

The flow rate, mean age, and drainage area all range several orders of magnitude and are log-transformed before spatial scaling analysis. The plot of log-transformed flow rate, log-transformed mean age, and solute concentration versus log-transformed drainage area shows significant scatter (not shown). The scatter can be attributed to the complexity of the 3-D topography and topographically driven groundwater flow system (Cardenas, 2007; Tóth, 1963). In order to eliminate the influence of scatter on spatial scaling analysis, the log-transformed flow rate, log-transformed mean age, and solute concentration are binned into intervals of log-transformed drainage area, following the approach in Kirchner (2009), and the median value and IQRs within each bin are calculated and plotted, with linear trends fitted to the median values (Figures 9 and S5).

Figures 9 and S5 show strong scaling relationships between flow rate, mean age, solute concentration, and drainage area. For models with deep extinction depths ($d_e \geq 33.3$ m), the spatial scaling of baseflow mean age and solute concentration is characterized by a positive slope (Figures 9h–9l, 9o-9s, S5h–S5l, and S5o–S5s). This scaling pattern qualitatively matches the observed spatial scaling of solute concentration from watersheds with significant deep groundwater contribution to streamflow (e.g., Figures 7 and 8 in Frisbee et al., 2017) and the 3-D catchment-mixing streamflow generation conceptual model (see Figure 1 in Frisbee et al., 2011). For models with shallow extinction depths ($d_e \leq 10$ m), the spatial scaling of baseflow mean age and solute concentration shows asymptotic behavior (Figures 9m and 9n, 9t and 9u, S5m and S5n, and S5t and S5u) and is consistent with observations in watersheds with dominant shallow flow contribution to streamflow (e.g., Figure 2 in Temnerud & Bishop, 2005) and 2-D network-mixing streamflow generation conceptual model (Figure 1 in Frisbee et al., 2011).

Moreover, although linear trends are fitted to all the plots in Figures 9 and S5, close inspection of the spatial scaling of baseflow age and solute concentration from models with deep extinction depths (Figures 9h–9l, 9o–9s, S5h–S5l, and S5o–S5s) reveals that, as drainage area increases, the mean age and solute concentration decreases when drainage area is smaller than a threshold drainage area of about 1×10^5 m^2 and increases at larger scales. This check mark shape scaling pattern is also observed in a real watershed (Figures 6 and 9 in Frisbee et al., 2011). However, Frisbee et al. (2011) were not able to give solid explanation to these distinct trends at two different scales due to the limited amount of field observations and hypothesized that it might be caused by the influence of a caldera wall at the headwaters of the studied volcanic watershed on the groundwater flow path (Frisbee et al., 2011). Interestingly, our models with simple geological heterogeneities but different extinction depths ($d_e \geq 33.3$ m) all produce the two trends at two scales. This indicates that the two trends at two scales are probably also closely related to the 3-D topographic structure and the complex topographically driven groundwater flow system, in addition to the geological structures such as a caldera wall.

The flow rate more closely follow its fitted linear trend than do the mean age and the solute concentration (Figures 9 and S5). The effect of mesh TCL and geological heterogeneity on the slopes of scaling between the flow rate and drainage area is indistinguishable within their confidence intervals (Figures 9a–9g, S5a–S5g, and S6a–S6g). Thus, the spatial scaling between baseflow and drainage area is not sensitive to capturing local-scale topographic variability along streams or ridges. This is partly because the flow rate cannot distinguish water from different flow paths as long as the aggregated amount of water into each stream reach is not significantly affected and thus cannot fully reveal the change of internal flow and transport characteristics due to the change of mesh TCL.

For baseflow mean age (Figures 9h–9n and S5h–S5n) and solute concentration (Figures 9o–9u and S5o–S5u), the binned values are more scattered around the fitted trends than is the flow rate. This indicates that the transport process and stream water chemistry are affected more by multiscale topographic variability and biogeochemical processes, which is also supported by field observations (e.g., McGuire et al., 2014).

The change of scaling slope of mean age due to mesh TCL or geological heterogeneity is not significant when extinction depth $d_e \geq 33.3$ m (Figures 9h–9l, S5h–S5l, and S6h–S6l). But the scaling slope decreases as the extinction depth further decreases (Figures 9m and 9n, S5m and S5n, and S6m and S6n). When $d_e \leq 10$ m, the probability density function of baseflow age has a heavy tail (e.g., Cardenas & Jiang, 2010). This is because although the regional-scale flow paths are suppressed in models with shallow extinction depth, the very old water from the part of the modeling domain with low hydraulic conductivity, although of low probability density, still contributes to a large proportion of the baseflow draining out of the domain. Since the mean value can be influenced by the heavy tail of probability density function, baseflow mean age is very old and relatively uniform across the whole watershed (Figures 9m and 9n, S5m and S5n, and S6m and S6n).

When the extinction depth is extremely shallow (d_e =3.3 m), the scaling slope between baseflow mean age and drainage area becomes negative. This is because the groundwater flow system is dominated by local flow paths when d_e is extremely small (e.g., Figure 3 in Cardenas & Jiang, 2010). The topography is more rugged in the upstream area and results in longer and deeper local flow paths than in the downstream area. Groundwater from longer and deeper local flow systems has higher mean age, and thus, the baseflow mean age decreases from upstream to downstream. Hence as we capture more low-order streams or ridges in the model (from TCL1S to TCL4S and from TCL1SR to TCL4SR), the local flow paths in the upstream area get longer and deeper and generate older baseflow than in the downstream area, which results in a decrease of the scaling slope as TCL increases (Figures 9n, S5n, and S6n). The change of spatial scaling pattern of solute concentration is similar to that of mean age (Figures 9o–9u, S5o–S5u, and S6o–S6u).

The low sensitivity of these trends to capturing local topographic variability when d_e is relatively large indicates that these trends can be general in regions dominated by topographically driven groundwater flow system, despite the differences in local-scale topographic variability. But the spatial scaling can be more sensitive to capturing local-scale topographic variability in systems with very shallow extinction depth.

3.4. Discussion of Specified Water Table Assumptions and Model Configuration

In this section, we explain why the Tothian assumption, where the water table mimics the topography, is suitable and necessary for the purposes of this study. We also discuss the appropriateness of the idealized model configuration.

3.4.1. Topography as Specified Water Table Assumption

Assuming that the water table is a replica of the land surface topography is appropriate for our research purpose for three main reasons: First, this study is theoretical in nature, and we are not aiming at reproducing the flow and transport pattern for any specific research site but to gain mechanistic insight. So this assumption does not raise an issue of uncertainty relative to specific site observations (Bresciani et al., 2016). Using land surface elevation as specified head top boundary condition avoids the uncertainty and complexity of specifying actual water table or introducing recharge and drain top boundary conditions and enables us to focus on the effect of capturing ridge and stream topographic features on the simulated flow and transport pattern. Second, this assumption avoids the need to solve the highly nonlinear unsaturated subsurface flow and dynamic water table, allowing us to explore a myriad of scenarios. A maximum node spacing of 80 m, which is equal to half the average divide-to-stream distance of the watershed, is used to ensure that the shape of the hillslopes can be captured. Further refining the mesh along selected streams and ridges results in 3,101,640 nodes for topography scenario TCL1S and 14,837,010 nodes for topography scenario TCL4SR. A parsimonious configuration for the flow process and flow boundary condition is critical for the implementation of this

study because the large number of numerical mesh nodes used to capture the complexity of the topography makes solving the unsaturated-saturated flow problem practically infeasible. And third, we believe our conclusions should raise attention to the research community on the importance of topography variability along lower-order streams, which are typically ignored in watershed scale groundwater modeling. This source of uncertainty should be evaluated whenever it is possible for simple model configurations like the one used in our study or for more complex fully coupled surface water-groundwater flow and transport models.

3.4.2. Appropriateness of Model Configuration

Ideally, to corroborate the conclusion with observations, models with meshes of different TCLs should be calibrated to match the observed streamflow, stream solute concentration, etc., and then the performance of models with different meshes should be evaluated. To do this, a fully integrated surface water-unsaturated-saturated groundwater model will be a better choice than the simple model configuration currently used in order to reproduce the realistic hydrologic processes in the watershed. However, as stated in section 3.4.1, the spirit of this work and the large number of nodes used to capture topographic variability prohibits us from further including complex flow processes and boundary conditions in the model. Besides this, detailed hydrogeologic conceptualization will also be needed. This will, on one hand, call for more geologic survey data, which are not available and, on the other hand, constrain our results to a single watershed and prohibit us from using various hydraulic conductivity scenarios to generalize our analysis to watersheds with different bedrock permeability conditions.

Another issue regarding corroborating the conclusion with observations is whether we want to use observations on transient hydrologic processes or observations representing steady state hydrologic conditions. Although some spatially distributed observations of the Rio Hondo baseflow and solute concentrations were collected by our colleagues (Frisbee et al., 2017), they are not long enough to be used to calibrate/validate a transient model or to represent the steady state hydrologic conditions in the watershed. Spatially distributed baseflow chemistry measurements are also available from studies on some other watersheds (e.g., Manning, 2011; Manning et al., 2012; Rademacher et al., 2001, 2005), but they are not comprehensive enough to corroborate our model results.

Despite of the infeasibility of building fully integrated models to represent realistic hydrologic processes and the lack of comprehensive observations to validate the modeling results, our simple model configuration with topography-driven groundwater flow assumption and hydraulic conductivity profiles of different extinction depths qualitatively reproduces various hydrologic features observed in natural watersheds. For example, the spatial scaling pattern of solute concentration from models including deep groundwater flow (i.e., models with high extinction depths) is consistent with the 3-D catchment-mixing streamflow generation conceptual model that has been tested with field observations in the Saguache and Rio Hondo watersheds (Frisbee et al., 2011, 2017). Similarly, the spatial scaling pattern of solute concentration for models with shallow extinction depths is consistent with the 2-D network-mixing streamflow generation conceptual model that has been tested in other watersheds (e.g., Temnerud & Bishop, 2005; Uchida et al., 2005). These provide evidence that the model configuration used in this study is appropriate while using Rio Hondo topography to test our hypothesis.

Building a realistic model for the Rio Hondo watershed or other watersheds to reproduce the realistic hydrologic processes is not the main purpose of this study. Instead, our focus is on the biasing effect caused by not capturing ridges and lower-order streams in the numerical mesh for the physics-based distributed groundwater flow and transport models. We simply use the topography of the Rio Hondo, a typical mountainous watershed, to illustrate the importance of capturing streams and ridges. We believe that the biasing effect revealed by our analysis exists in the modeling results whether the model is corroborated by the observations or not, and this kind of uncertainty should be considered while modeling actual hydrologic processes in natural watersheds.

4. Conclusions

Using various hydrologic metrics, we study the effect of capturing streams and ridges with different resolutions during mesh generation. Eight mesh scenarios with increasing topographic complexity level (TCL) are created by progressively refining the mesh along low-order streams and ridges. We also explore the influence that systematic changes in geological heterogeneity have on the importance of mesh TCL by using one homogeneous hydraulic conductivity and six different exponentially depth-decaying hydraulic conductivity

fields. The seven hydraulic conductivity fields and eight mesh TCLs result in a total of 56 modeling scenarios. For all the scenarios, we simulate steady state saturated three-dimensional groundwater flow and mean age. Multiple lumped and distributed metrics including net baseflow generated within the whole domain, flushing intensity of groundwater flow, baseflow into stream reaches of different orders, flux-weighted baseflow mean age, baseflow solute concentration from subsurface chemical weathering, and spatial scaling relationship between hydrological variables and drainage area are calculated to evaluate the effect of mesh TCL under different geological conditions. Our results imply that capturing streams and ridges is fundamental to reproduce realistic internal groundwater flow and transport characteristics, in addition to the integrated response like the hydrograph at the watershed outlet (Kirchner, 2006).

The main conclusions are as follows:

Mesh TCL has a relatively small influence on the integrated metrics including net flux, flushing intensity, and spatial scaling of flow rate, mean age, and solute concentration. However, mesh TCL can significantly affect the relative amount of groundwater through subsurface flow paths of different scales and the relative amount of baseflow into streams of different orders. When compared to the model with the highest TCL, lower TCL models capturing less topographic features, such as low-order streams or ridges, produce less flow through short, local-scale flow paths and more flow through longer, intermediate and regional-scale flow paths. This results in less baseflow into low-order streams and more baseflow into high-order streams. The baseflow rate bias due to missing streams or ridges can be up to 10% in systems with deep extinction depths of hydraulic conductivity fields and 20% in systems with shallow extinction depths of hydraulic conductivity fields.

Meshes that ignore low-order streams and ridges can reduce flow through shallow and short local-scale flow paths and increase flow through deep and long regional-scale flow paths. This results in lower amount of young water flowing into low-order streams and larger amount of old water flowing into high-order streams. The overall effect is that the baseflow into streams of all orders is biased older in models with low TCL than the models with high TCL which capture more stream or ridge features.

Baseflow solute concentration is less sensitive to model TCL due to the constraint of thermodynamic limit on equilibrium concentration. In models with rapidly decaying hydraulic conductivity or for high-order streams, the mean age of groundwater discharging into streams is longer than equilibrium time of chemical reaction. So in these cases the solute concentration approaches equilibrium concentration and is affected little by mesh TCL. For low-order streams in models with slowly decaying hydraulic conductivity, the mean age of groundwater discharging into streams is shorter than chemical equilibrium time and the change of mean age caused by mesh TCL can affect solute concentration in these low-order streams.

A shallower extinction depth of hydraulic conductivity enhances the effect of mesh TCL on flow rate and mean age of baseflow, because the subsurface flow system is forced at the upper part of the modeling domain and is affected more by capturing local water table variability. The effect of mesh TCL should be more significant in models of regions with shallow and low permeability bedrock.

We also suspect that the effect of mesh TCL can be more prominent in models of rugged watersheds with deep cutting streams and valleys, and high and steep hillslopes, than in relatively flat watersheds. Rugged and steep watersheds are typical in mountainous regions, which serves as the water source for mountain front floodplain and farmland. Thus, neglecting the topographic features of streams and ridges in model mesh generation can seriously bias the simulated spatial distribution of streamflow generation, riparian vegetation, and evapotranspiration and affect water resource management. The bias caused by not refining the mesh along low-order streams should be taken into consideration while interpreting modeling results, especially in mountainous watersheds with complex topographic variability. Future research can relax the Tothian assumption and use more realistic boundary conditions at the surface by incorporating unsaturated zone processes.

Acknowledgments
This work was supported through funding from the New Mexico Water Resources Research Institute Faculty Water Research Grant, the National Science Foundation (EAR-1015100, EAR-1830172, and CNH-1010516), and the New Mexico EPSCoR Track I (EAR-0814449) awarded to New Mexico Tech. Wang and Gomez-Velez are also funded by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of BER's Subsurface Biogeochemistry Research Program (SBR). This contribution originates from the SBR Scientific Focus Area (SFA) at the Pacific Northwest National Laboratory (PNNL). We thank ARANZ Geo Limited for kindly providing the Leapfrog academic licenses used to generate the modeling meshes. Mention of trade names or commercial products does not constitute endorsement or recommendation for use. The work is theoretical in nature. It does not use observational data, and the digital elevation model for the Rio Hondo watershed is publicly available from the USGS data portal.

References

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

Bao, C., Li, L., Shi, Y., & Duffy, C. (2017). Understanding watershed hydrogeochemistry: 1. Development of RT-Flux-PIHM. *Water Resources Research*, *53*, 2328–2345. https://doi.org/10.1002/2016WR018934

Bear, J. (1972). *Dynamics of fluids in porous media*. New York: Elsevier.

Beven, K. (1989). Changing ideas in hydrology—The case of physically-based models. *Journal of Hydrology*, *105*(1), 157–172. https://doi.org/10.1016/0022-1694(89)90101-7

Beven, K. (2001). How far can we go in distributed hydrological modelling? *Hydrology and Earth System Sciences*, *5*(1), 1–12. https://doi.org/10.5194/hess-5-1-2001

Beven, K., & Freer, J. (2001). A dynamic TOPMODEL. *Hydrological Processes*, *15*(10), 1993–2011. https://doi.org/10.1002/hyp.252

Beven, K. J., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, *24*(1), 43–69. https://doi.org/10.1080/02626667909491834

Boano, F., Harvey, J. W., Marion, A., Packman, A. I., Revelli, R., Ridolfi, L., & Wörman, A. (2014). Hyporheic flow and transport processes: Mechanisms, models, and biogeochemical implications. *Reviews of Geophysics*, *52*, 603–679. https://doi.org/10.1002/2012RG000417

Bresciani, E., Gleeson, T., Goderniaux, P., Dreuzy, J. R., Werner, A. D., Wörman, A., et al. (2016). Groundwater flow systems theory: Research challenges beyond the specified-head top boundary condition. *Hydrogeology Journal*, *24*, 1–4. https://doi.org/10.1007/s10040-016-1397-8

Buffington, J. M., & Tonina, D. (2009). Hyporheic exchange in mountain rivers II: Effects of channel morphology on mechanics, scales, and rates of exchange. *Geography Compass*, *3*(3), 1038–1062. https://doi.org/10.1111/j.1749-8198.2009.00225.x

Cardenas, M. B. (2007). Potential contribution of topography-driven regional groundwater flow to fractal stream chemistry: Residence time distribution analysis of Tóth flow. *Geophysical Research Letters*, *34*, L05403. https://doi.org/10.1029/2006GL029126

Cardenas, M. B. (2008). Surface water-groundwater interface geomorphology leads to scaling of residence times. *Geophysical Research Letters*, *35*, L08402. https://doi.org/10.1029/2008GL033753

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

Caruso, A., Ridolfi, L., & Boano, F. (2016). Impact of watershed topography on hyporheic exchange. *Advances in Water Resources*, *94*, 400–411. https://doi.org/10.1016/j.advwatres.2016.06.005

Chen, J. (2015). Impact of climate change on Canadian water resources: A continental-scale hydrologic modelling study using multiple RCM projections (PhD Dissertation), Waterloo, Ontario, Canada.

Chow, R., Wu, H., Bennett, J. P., Dugge, J., Wöhling, T., & Nowak, W. (2018). Sensitivity of simulated hyporheic exchange to river bathymetry: The Steinlach River Test Site. *Groundwater*. https://doi.org/10.1111/gwat.12816

Clark, M. P., Kavetski, D., & Fenicia, F. (2011). Pursuing the method of multiple working hypotheses for hydrological modeling. *Water Resources Research*, *47*, W09301. https://doi.org/10.1029/2010WR009827

Condon, L. E., & Maxwell, R. M. (2015). Evaluating the relationship between topography and groundwater using outputs from a continental-scale integrated hydrology model. *Water Resources Research*, *51*, 6602–6621. https://doi.org/10.1002/2014WR016774

Dingman, S. L. (2015). *Physical hydrology* (3rd ed.). IL, USA: Waveland Press.

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

Frei, S., Lischeid, G., & Fleckenstein, J. H. (2010). Effects of micro-topography on surface–subsurface exchange and runoff generation in a virtual riparian wetland — A modeling study. *Advances in Water Resources*, *33*(11), 1388–1401. https://doi.org/10.1016/j.advwatres.2010.07.006

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

Gelhar, L. W., Welty, C., & Rehfeldt, K. R. (1992). A critical review of data on field-scale dispersion in aquifers. *Water Resources Research*, *28*(7), 1955–1974. https://doi.org/10.1029/92WR00607

Genereux, D. P., Nagy, L. A., Osburn, C. L., & Oberbauer, S. F. (2013). A connection to deep groundwater alters ecosystem carbon fluxes and budgets: Example from a Costa Rican rainforest. *Geophysical Research Letters*, *40*, 2066–2070. https://doi.org/10.1002/grl.50423

Ginn, T. R. (1999). On the distribution of multicomponent mixtures over generalized exposure time in subsurface flow and reactive transport: Foundations, and formulations for groundwater age, chemical heterogeneity, and biodegradation. *Water Resources Research*, *35*(5), 1395–1407. https://doi.org/10.1029/1999WR900013

Glaser, B., Klaus, J., Frei, S., Frentress, J., Pfister, L., & Hopp, L. (2016). On the value of surface saturated area dynamics mapped with thermal infrared imagery for modeling the hillslope-riparian-stream continuum. *Water Resources Research*, *52*, 8317–8342. https://doi.org/10.1002/2015WR018414

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

Gleeson, T., Marklund, L., Smith, L., & Manning, A. H. (2011). Classifying the water table at regional to continental scales. *Geophysical Research Letters*, *38*, L05401. https://doi.org/10.1029/2010GL046427

Goderniaux, P., Brouyère, S., Wildemeersch, S., Therrien, R., & Dassargues, A. (2015). Uncertainty of climate change impact on groundwater reserves—Application to a chalk aquifer. *Journal of Hydrology*, *528*, 108–121. https://doi.org/10.1016/j.jhydrol.2015.06.018

Gomez, J. D., & Wilson, J. L. (2013). Age distributions and dynamically changing hydrologic systems: Exploring topography-driven flow. *Water Resources Research*, *49*, 1503–1522. https://doi.org/10.1002/wrcr.20127

Gomez-Velez, J. D., Krause, S., & Wilson, J. L. (2014). Effect of low-permeability layers on spatial patterns of hyporheic exchange and groundwater upwelling. *Water Resources Research*, *50*, 5196–5215. https://doi.org/10.1002/2013WR015054

Goode, D. J. (1996). Direct simulation of groundwater age. *Water Resources Research*, *32*(2), 289–296.

Green, M. B., & Wang, D. (2008). Watershed flow paths and stream water nitrogen-to-phosphorus ratios under simulated precipitation regimes. *Water Resources Research*, *44*, W12414. https://doi.org/10.1029/2007WR006139

Güntner, A., Uhlenbrook, S., Seibert, J., & Leibundgut, Ch. (1999). Multi-criterial validation of TOPMODEL in a mountainous catchment. *Hydrological Processes*, *13*(11), 1603–1620. https://doi.org/10.1002/(SICI)1099-1085(19990815)13:11<1603::AID-HYP830>3.3.CO;2-B

Gupta, H. V., Clark, M. P., Vrugt, J. A., Abramowitz, G., & Ye, M. (2012). Towards a comprehensive assessment of model structural adequacy. *Water Resources Research*, *48*, W08301. https://doi.org/10.1029/2011WR011044

Gupta, H. V., & Nearing, G. S. (2014). Debates—the future of hydrological sciences: A (common) path forward? Using models and data to learn: A systems theoretic perspective on the future of hydrological science. *Water Resources Research*, *50*, 5351–5359. https://doi.org/10.1002/2013WR015096

Haitjema, H. M., & Mitchell-Bruker, S. (2005). Are water tables a subdued replica of the topography? *Ground Water*, *43*(6), 781–786. 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

Hartmann, A., Gleeson, T., Wada, Y., & Wagener, T. (2017). Enhanced groundwater recharge rates and altered recharge sensitivity to climate variability through subsurface heterogeneity. *Proceedings of the National Academy of Sciences*, *114*(11), 2842–2847. https://doi.org/10.1073/pnas.1614941114

Healy, R. W., & Scanlon, B. R. (2010). *Estimating groundwater recharge*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511780745

Horritt, M. S., Bates, P. D., & Mattinson, M. J. (2006). Effects of mesh resolution and topographic representation in 2D finite volume models of shallow water fluvial flow. *Journal of Hydrology*, *329*(1), 306–314. https://doi.org/10.1016/j.jhydrol.2006.02.016

Horton, R. E. (1932). Drainage-basin characteristics. *Transactions, American Geophysical Union*, *13*(1), 350–361. https://doi.org/10.1029/TR013i001p00350

Houser, P. R., Shuttleworth, W. J., Famiglietti, J. S., Gupta, H. V., Syed, K. H., & Goodrich, D. C. (1998). Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. *Water Resources Research*, *34*(12), 3405–3420. https://doi.org/10.1029/1998WR900001

Ingebritsen, S. E., & Manning, C. E. (1999). Geological implications of a permeability-depth curve for the continental crust. *Geology*, *27*(12), 1107–1110. https://doi.org/10.1130/0091-7613(1999)027<1107:GIOAPD>2.3.CO;2

Kirchner, J. W. (2006). Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research*, *42*, W03S04. https://doi.org/10.1029/2005WR004362

Kirchner, J. W. (2009). Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resources Research*, *45*, W02429. https://doi.org/10.1029/2008WR006912

Kirchner, J. W., Feng, X., & Neal, C. (2000). Fractal stream chemistry and its implications for contaminant transport in catchments. *Nature*, *403*(6769), 524–527. https://doi.org/10.1038/35000537

Kirchner, J. W., & Neal, C. (2013). Universal fractal scaling in stream chemistry and its implications for solute transport and water quality trend detection. *Proceedings of the National Academy of Sciences*, *110*(30), 12,213–12,218. https://doi.org/10.1073/pnas.1304328110

Kollet, S. J., & Maxwell, R. M. (2008). Demonstrating fractal scaling of baseflow residence time distributions using a fully-coupled groundwater and land surface model. *Geophysical Research Letters*, *35*, L07402. https://doi.org/10.1029/2008GL033215

Kuo, W.-L., Steenhuis, T. S., McCulloch, C. E., Mohler, C. L., Weinstein, D. A., DeGloria, S. D., & Swaney, D. P. (1999). Effect of grid size on runoff and soil moisture for a variable-source-area hydrology model. *Water Resources Research*, *35*(11), 3419–3428. https://doi.org/10.1029/1999WR900183

Li, L., Bao, C., Sullivan., P. L., Brantley, S., Shi, Y., & Duffy, C. (2017). Understanding watershed hydrogeochemistry: 2. Synchronized hydrological and geochemical processes drive stream chemostatic behavior. *Water Resources Research*, *53*, 2346–2367. https://doi.org/10.1002/2016WR018935

Li, H., & Sivapalan, M. (2011). Effect of spatial heterogeneity of runoff generation mechanisms on the scaling behavior of event runoff responses in a natural river basin. *Water Resources Research*, *47*, W00H08. https://doi.org/10.1029/2010WR009712

Li, H., Wigmosta, M. S., Wu, H., Huang, M., Ke, Y., Coleman, A. M., & Leung, L. R. (2013). A physically based runoff routing model for land surface and Earth system models. *Journal of Hydrometeorology*, *14*(3), 808–828. https://doi.org/10.1175/JHM-D-12-015.1

Liu, Y., Weerts, A. H., Clark, M., Hendricks Franssen, H.-J., Kumar, S., Moradkhani, H., et al. (2012). Advancing data assimilation in operational hydrologic forecasting: Progresses, challenges, and emerging opportunities. *Hydrology and Earth System Sciences*, *16*(10), 3863–3887. https://doi.org/10.5194/hess-16-3863-2012

Louis, C. (1972). Rock hydraulics. In L. Müller (Ed.), *Rock mechanics, International Centre for Mechanical Sciences* (pp. 299–387). Vienna: Springer. https://doi.org/10.1007/978-3-7091-4109-0_16

Maher, K. (2011). The role of fluid residence time and topographic scales in determining chemical fluxes from landscapes. *Earth and Planetary Science Letters*, *312*(1), 48–58. https://doi.org/10.1016/j.epsl.2011.09.040

Maher, K., & Chamberlain, C. P. (2014). Hydrologic regulation of chemical weathering and the geologic carbon cycle. *Science*, *343*(6178), 1502–1504. https://doi.org/10.1126/science.1250770

Mahmood, T. H., & Vivoni, E. R. (2011). Breakdown of hydrologic patterns upon model coarsening at hillslope scales and implications for experimental design. *Journal of Hydrology*, *411*(3), 309–321. https://doi.org/10.1016/j.jhydrol.2011.10.011

Manning, A. H. (2011). Mountain-block recharge, present and past, in the eastern Espanola Basin, New Mexico, USA. *Hydrogeology Journal*, *19*(2), 379–397.

Manning, A. H., Clark, J. F., Diaz, S. H., Rademacher, L. K., Earman, S., & Niel Plummer, L. (2012). Evolution of groundwater age in a mountain watershed over a period of thirteen years. *Journal of Hydrology*, *460*, 13–28. https://doi.org/10.1016/j.jhydrol.2012.06.030

Marklund, L., & Wörman, A. (2011). The use of spectral analysis-based exact solutions to characterize topography-controlled groundwater flow. *Hydrogeology Journal*, *19*(8), 1531–1543. https://doi.org/10.1007/s10040-011-0768-4

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

McDonnell, J. J., Sivapalan, M., Vaché, K., Dunn, S., Grant, G., Haggerty, R., et al. (2007). Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. *Water Resources Research*, *43*, W07301. https://doi.org/10.1029/2006WR005467

McGuire, K. J., McDonnell, J. J., Weiler, M., Kendall, C., McGlynn, B. L., Welker, J. M., & Seibert, J. (2005). The role of topography on catchment-scale water residence time. *Water Resources Research*, *41*, W05002. https://doi.org/10.1029/2004WR003657

McGuire, K. J., Torgersen, C. E., Likens, G. E., Buso, D. C., Lowe, W. H., & Bailey, S. W. (2014). Network analysis reveals multiscale controls on streamwater chemistry. *Proceedings of the National Academy of Sciences*, *111*(19), 7030–7035. https://doi.org/10.1073/pnas.1404820111

O'Loughlin, E. M. (1986). Prediction of surface saturation zones in natural catchments by topographic analysis. *Water Resources Research*, *22*(5), 794–804. https://doi.org/10.1029/WR022i005p00794

Park, S. K., & Xu, L. (Eds.) (2009). *Data assimilation for atmospheric, oceanic and hydrologic applications*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Partington, D., Brunner, P., Frei, S., Simmons, C. T., Werner, A. D., Therrien, R., et al. (2013). Interpreting streamflow generation mechanisms from integrated surface-subsurface flow models of a riparian wetland and catchment. *Water Resources Research*, *49*, 5501–5519. https://doi.org/10.1002/wrcr.20405

Peralta-Tapia, A., Sponseller, R. A., Ågren, A., Tetzlaff, D., Soulsby, C., & Laudon, H. (2015). Scale-dependent groundwater contributions influence patterns of winter baseflow stream chemistry in boreal catchments. *Journal of Geophysical Research: Biogeosciences*, *120*, 847–858. https://doi.org/10.1002/2014JG002878

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

Rademacher, L. K., Clark, J. F., Hudson, G. B., Erman, D. C., & Erman, N. A. (2001). Chemical evolution of shallow groundwater as recorded by springs, Sagehen basin; Nevada County, California. *Chemical Geology*, *179*(1), 37–51. https://doi.org/10.1016/S0009-2541(01)00314-X

Riml, J., & Wörman, A. (2011). Response functions for in-stream solute transport in river networks. *Water Resources Research*, *47*, W06502. https://doi.org/10.1029/2010WR009412

Rodríguez-Iturbe, I., & Rinaldo, A. (1997). *Fractal river basins: Chance and self-organization*. Cambridge: Cambridge University Press.

Saar, M. O., & Manga, M. (2004). Depth dependence of permeability in the Oregon Cascades inferred from hydrogeologic, thermal, seismic, and magmatic modeling constraints. *Journal of Geophysical Research*, *109*, B04204. https://doi.org/10.1029/2003JB002855

Santhi, C., Allen, P. M., Muttiah, R. S., Arnold, J. G., & Tuppad, P. (2008). Regional estimation of base flow for the conterminous United States by hydrologic landscape regions. *Journal of Hydrology*, *351*(1), 139–153. https://doi.org/10.1016/j.jhydrol.2007.12.018

Scanlon, B. R., Healy, R. W., & Cook, P. G. (2002). Choosing appropriate techniques for quantifying groundwater recharge. *Hydrogeology Journal*, *10*(1), 18–39. https://doi.org/10.1007/s10040-001-0176-2

Shen, C., Riley, W. J., Smithgall, K. R., Melack, J. M., & Fang, K. (2016). The fan of influence of streams and channel feedbacks to simulated land surface water and carbon dynamics. *Water Resources Research*, *52*, 880–902. https://doi.org/10.1002/2015WR018086

Singh, N. K., Emanuel, R. E., & McGlynn, B. L. (2016). Variability in isotopic composition of base flow in two headwater streams of the southern Appalachians. *Water Resources Research*, *52*, 4264–4279. https://doi.org/10.1002/2015WR018463

Singh, R. S., Reager, J. T., Miller, N. L., & Famiglietti, J. S. (2015). Toward hyper-resolution land-surface modeling: The effects of fine-scale topography and soil texture on CLM4.0 simulations over the Southwestern U.S. *Water Resources Research*, *51*, 2648–2667. https://doi.org/10.1002/2014WR015686

Sivapalan, M. (2003). Process complexity at hillslope scale, process simplicity at the watershed scale: Is there a connection? *Hydrological Processes*, *17*(5), 1037–1041. https://doi.org/10.1002/hyp.5109

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., et al. (2003). IAHS decade on predictions in ungauged basins (PUB), Shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal*, *48*(6), 857–880. https://doi.org/10.1623/hysj.48.6.857.51421

Stocker, B. D., Spahni, R., & Joos, F. (2014). DYPTOP: A cost-efficient TOPMODEL implementation to simulate sub-grid spatio-temporal dynamics of global wetlands and peatlands. *Geoscientific Model Development*, *7*(6), 3089–3110. https://doi.org/10.5194/gmd-7-3089-2014

Stoertz, M. W., & Bradbury, K. R. (1989). Mapping recharge areas using a ground-water flow model—A case study. *Groundwater*, *27*(2), 220–228. https://doi.org/10.1111/j.1745-6584.1989.tb00443.x

Su, C.-H., Peterson, T. J., Costelloe, J. F., & Western, A. W. (2016). A synthetic study to evaluate the utility of hydrological signatures for calibrating a base flow separation filter. *Water Resources Research*, *52*, 6526–6540. https://doi.org/10.1002/2015WR018177

Sudicky, E. A., Jones, J. P., Park, Y.-J., Brookfield, A. E., & Colautti, D. (2008). Simulating complex flow and transport dynamics in an integrated surface-subsurface modeling framework. *Geosciences Journal*, *12*(2), 107–122. https://doi.org/10.1007/s12303-008-0013-x

Sulis, M., Paniconi, C., & Camporese, M. (2011). Impact of grid resolution on the integrated and distributed response of a coupled surface–subsurface hydrological model for the des Anglais catchment, Quebec. *Hydrological Processes*, *25*(12), 1853–1865. https://doi.org/10.1002/hyp.7941

Tarboton, D. G., Bras, R. L., & Rodriguez-Iturbe, I. (1991). On the extraction of channel networks from digital elevation data. *Hydrological Processes*, *5*(1), 81–100. https://doi.org/10.1002/hyp.3360050107

Temnerud, J., & Bishop, K. (2005). Spatial variation of streamwater chemistry in two Swedish boreal catchments: Implications for environmental assessment. *Environmental Science & Technology*, *39*(6), 1463–1469. https://doi.org/10.1021/es040045q

Therrien, R., McLaren, R. G., Sudicky, E. A., & Panday, S. M. (2010). HydroGeoSphere: A three-dimensional numerical model describing fully-integrated subsurface and surface flow and solute transport. Groundwater Simulations Group, University of Waterloo, Waterloo, ON.

Tolley, D. G. (2014). High-elevation mountain streamflow generation: The role of deep groundwater in the Rio Hondo Watershed, Northern New Mexico (MS Thesis), New Mexico Institute of Mining and Technology, Socorro, New Mexico, USA.

Tolley, D. G., Frisbee, M. D., & Campbell, A. R. (2015). Determining the importance of seasonality on groundwater recharge and streamflow in the Sangre de Cristo Mountains using stable isotopes. In J. Lindline, et al. (Eds.), *Geology of the Las Vegas region* (vol. Guidebook, 66th Field Conference, pp. 303–312). Socorro, New Mexico: New Mexico Geological Society.

Tonina, D., & Buffington, J. M. (2009). Hyporheic exchange in mountain rivers I: Mechanics and environmental effects. *Geography Compass*, *3*(3), 1063–1086. https://doi.org/10.1111/j.1749-8198.2009.00226.x

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

Tóth, J. (2009). *Gravitational systems of groundwater flow: Theory, evaluation, utilization*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511576546

Uchida, T., Asano, Y., Onda, Y., & Miyata, S. (2005). Are headwaters just the sum of hillslopes? *Hydrological Processes*, *19*(16), 3251–3261. https://doi.org/10.1002/hyp.6004

Varni, M., & Carrera, J. (1998). Simulation of groundwater age distributions. *Water Resources Research*, *34*(12), 3271–3281.

Vieux, B. E. (1993). DEM aggregation and smoothing effects on surface runoff modeling. *Journal of Computing in Civil Engineering*, *7*(3), 310–338. https://doi.org/10.1061/(ASCE)0887-3801(1993)7:3(310)

Vivoni, E. R., Ivanov, V. Y., Bras, R. L., & Entekhabi, D. (2004). Generation of triangulated irregular networks based on hydrological similarity. *Journal of Hydrologic Engineering*, *9*(4), 288–302. https://doi.org/10.1061/(ASCE)1084-0699(2004)9:4(288)

Vivoni, E. R., Ivanov, V. Y., Bras, R. L., & Entekhabi, D. (2005). On the effects of triangulated terrain resolution on distributed hydrologic model response. *Hydrological Processes*, *19*(11), 2101–2122. https://doi.org/10.1002/hyp.5671

Welch, L. A., Allen, D. M., & (Ilja) van Meerveld, H. J. (2012). Topographic controls on deep groundwater contributions to mountain headwater streams and sensitivity to available recharge. *Canadian Water Resources Journal*, *37*(4), 349–371. https://doi.org/10.4296/cwrj2011-907

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

Wilcox, B. P., Wilding, L. P., & Woodruff, C. M. (2007). Soil and topographic controls on runoff generation from stepped landforms in the Edwards Plateau of Central Texas. *Geophysical Research Letters*, *34*, L24S24. https://doi.org/10.1029/2007GL030860

Wilson, J. L., & Guan, H. (2004). Mountain-block hydrology and mountain-front recharge. In J. F. Hogan, et al. (Eds.), *Groundwater recharge in a desert environment: The southwestern United States* (Vol. 9, pp. 113–137). Washington, DC: American Geophysical Union. https://doi.org/10.1029/009WSA08

Winter, T. C. (1999). Relation of streams, lakes, and wetlands to groundwater flow systems. *Hydrogeology Journal*, *7*(1), 28–45. https://doi.org/10.1007/s100400050178

Winter, T. C., Rosenberry, D. O., & LaBaugh, J. W. (2003). Where does the ground water in small watersheds come from? *Groundwater*, *41*(7), 989–1000. https://doi.org/10.1111/j.1745-6584.2003.tb02440.x

Wolock, D. M., Hornberger, G. M., & Musgrove, T. J. (1990). Topographic effects on flow path and surface water chemistry of the Llyn Brianne catchments in Wales. *Journal of Hydrology*, *115*(1), 243–259. https://doi.org/10.1016/0022-1694(90)90207-E

Wörman, A, Packman, A. I., Marklund, L., Harvey, J. W., & Stone, S. H. (2007). Fractal topography and subsurface water flows from fluvial bedforms to the continental shield. *Geophysical Research Letters*, *34*, L07402. https://doi.org/10.1029/2007GL029426

Zhang, W., & Montgomery, D. R. (1994). Digital elevation model grid size, landscape representation, and hydrologic simulations. *Water Resources Research*, *30*(4), 1019–1028. https://doi.org/10.1029/93WR03553

Zijl, W. (1999). Scale aspects of groundwater flow and transport systems. *Hydrogeology Journal*, *7*(1), 139–150. https://doi.org/10.1007/s100400050185

Zlotnik, V. A., Cardenas, M. B., & Toundykov, D. (2011). Effects of multiscale anisotropy on basin and hyporheic groundwater flow. *Ground Water*, *49*(4), 576–583. https://doi.org/10.1111/j.1745-6584.2010.00775.x