Groundwater

Simulating Groundwater-Streamflow Connections in the Upper Colorado River Basin

by Hoang Tran[1], Jun Zhang[2], Jean-Martial Cohard[3], Laura E. Condon[2], and Reed M. Maxwell[4]

Abstract

In mountain, snow driven catchments, snowmelt is supposed to be the primary contribution to river streamflows during spring. In these catchments the contribution of groundwater is not well documented because of the difficulty to monitor groundwater in such complex environment with deep aquifers. In this study we use an integrated hydrologic model to conduct numerical experiments that help quantify the effect of lateral groundwater flow on total annual and peak streamflow in predevelopment conditions. Our simulations focus on the Upper Colorado River Basin (UCRB; 2.8×10^5 km^2) a well-documented mountain catchment for which both streamflow and water table measurements are available for several important sub-basins. For the simulated water year, our results suggest an increase in peak flow of up to 57% when lateral groundwater flow processes are included—an unexpected result for flood conditions generally assumed independent of groundwater. Additionally, inclusion of lateral groundwater flow moderately improved the model match to observations. The correlation coefficient for mean annual flows improved from 0.84 for the no lateral groundwater flow simulation to 0.98 for the lateral groundwater flow one. Spatially we see more pronounced differences between lateral and no lateral groundwater flow cases in areas of the domain with steeper topography. We also found distinct differences in the magnitude and spatial distribution of streamflow changes with and without lateral groundwater flow between Upper Colorado River Sub-basins. A sensitivity test that scaled hydraulic conductivity over two orders of magnitude was conducted for the lateral groundwater flow simulations. These results show that the impact of lateral groundwater flow is as large or larger than an order of magnitude change in hydraulic conductivity. While our results focus on the UCRB, we feel that these simulations have relevance to other headwaters systems worldwide.

Introduction

Mountain headwater basins can greatly affect downstream water deliveries (Alexander et al. 2007) controlling the flow of water to larger streams (Environmental Protection Agency 2015). Many headwater basins are located at high elevation and are snow-dominated which shifts water availability in time and then provides surface water during dry periods. Snow provides water for about one-sixth of the world's population and 70% of the Western United States' water supply (Chang et al. 1987; Barnett et al. 2005; Bales et al. 2006). The Upper Colorado River Basin (UCRB) system (2.8×10^5 km^2), one of the principal headwater basins in the United States, supplies water for over 40 million people, has several major dams, and also is the most overallocated river system in the world (Christensen et al. 2004).

Interannual variability in the Colorado River's flow can affect the water supply for millions of people. For example, the 15-year drought from 2000 to 2014, when streamflow reduced by around 30% (Woodhouse and Pederson 2018), caused serious shortages for water supplies in Arizona and Nevada. On the contrary, snowmelt following an anomalously wet winter in 1983 created the highest flows into Lake Powell on record, overwhelming the Glen Canyon Dam resulting in damage to the spillways (Vandivere and Vorster 1984).

Along with in situ observations and remote sensing data products, models are an important tool for understanding the dynamics of large complex, managed systems such as the UCRB and are the only tool for water forecasting. The Colorado river system has been modeled extensively using approaches that range in complexity (Tillman 2015). Examples include the SACramento soil moisture accounting model (SAC-SMA) (e.g., Nash and Gleick 1991; Franz et al. 2003), and

[1]Corresponding author: Integrated GroundWater Modeling Center and Department of Geology and Geological Engineering, Colorado School of Mines, Golden, CO; hoangtran@mines.edu
[2]Department of Hydrology and Atmospheric Sciences, The University of Arizona, Tucson, AZ; junzhang55@arizona.edu; lecondon@arizona.edu
[3]Univ. Grenoble Alpes, IRD, CNRS, Grenoble INP, IGE, 38000 Grenoble, France; jean-martial.cohard@univ-grenoble-alpes.fr
[4]Integrated GroundWater Modeling Center and Department of Geology and Geological Engineering, Colorado School of Mines, Golden, CO; rmaxwell@mines.edu

Received September 2019, accepted March 2020.

doi: 10.1111/gwat.13000

the variable infiltration capacity (VIC) model (Christensen et al. 2004; Christensen and Lettenmaier 2007; Painter et al. 2010). While both SAC-SMA and VIC are well established modeling platforms they simplify groundwater flow.

The role of lateral groundwater flow in watershed dynamics has been demonstrated in prior studies (e.g., Sear et al. 1999; Brouyère et al. 2004; Kollet and Maxwell 2008; Decharme et al. 2010; Flessa et al. 2013; Miller et al. 2016; Fang and Pomeroy 2016; Buto et al. 2017). On a large-scale, Condon and Maxwell (2019) and Condon et al. (2020) demonstrated the close connections between lateral groundwater flow streamflow and evapotranspiration (ET) using an integrated groundwater surface water model. However, few have investigated the role of lateral groundwater flow specifically in mountain headwater domains.

Here, we use an integrated hydrologic model, ParFlow-CLM (PF-CLM; Maxwell et al. 2015; Kollet et al. 2017), simulating the UCRB for predevelopment (i.e., no anthropogenic impacts such as reservoirs) in a historic flood year, a simulation time when groundwater is commonly assumed to play a more minor role in streamflow, to study the impact of including lateral groundwater flow on simulated streamflow and ET. Common to all integrated models is that they solve partial differential equations for surface (shallow water) and subsurface (Richards) flow in a globally implicit manner in three spatial dimensions. As such, these models are often more computationally expensive but represent many process connections, such as stream-aquifer interactions, implicitly and allow numerical experiments to be conducted with a range of physical parameterization to test model sensitivity.

In this study, we hypothesize that lateral groundwater flow has a significant effect on streamflow in a headwaters system even during wet years. We chose a high flow year when groundwater is generally thought to have a minor impact on streamflow. We perform a series of hypothetical numerical experiments with and without lateral groundwater flow to isolate the role of lateral groundwater flow under predevelopment conditions. These simulations use real topography, our best estimates of reconstructed hourly meteorology, land cover, soil and hydrostratigraphy with a-priori estimates of parameter values. Because saturated hydraulic conductivity is a sensitive, uncertain and important parameter that moderates groundwater flow, we conducted a sensitivity test, scaling these values over two orders of magnitude to help quantify the impact of this parameter on our conclusions. This paper is organized as follows: (1) we present the methods and the site; (2) we describe the numerical experiments performed; (3) we present our results focusing on differences between lateral groundwater flow and hydraulic conductivity scenarios; (4) we compare our model simulation results to observations to discuss sources of bias; and (5) we discuss potential impacts of lateral groundwater flow on water balance of sub regions within the UCRB.

Methods

Site

Our study focuses on the UCRB (Figure 1), an important headwater domain that encompasses 2.8×10^5 km^2 in the Rocky Mountains, including the headwaters of the Green, Yampa, Gunnison, Colorado and San Juan Rivers. Elevations in the UCRB ranges from peaks higher than 3300 m at the Rocky Mountains to around 900 m at Lee's Ferry (AZ), the outlet of the simulated hydrological domain (Figure 1). The basin is characterized by two different types of terrain. The Eastern parts of the basin are characterized by steep slopes and form the origins of the Gunnison, Colorado, and San Juan Rivers (Figure 1a). The Green River plateau has less topographic relief, with elevation ranging from 2000 to 2500 m. Rivers in this plateau, namely the Green and Yampa, are also generally flatter. The annual average temperature is $+6$ °C with average daily maximum temperatures ranging from -10 °C to $+14$ °C at low elevations (Kopytkovskiy et al. 2015), the UCRB is a snowmelt dominated system. Annual average precipitation for the Eastern part of the basin (the Rockies area) and the whole basin are 1000 and 164 mm, respectively. The annual average discharge at the outlet, Lee's Ferry, is approximately 420 m^3/s.

The UCRB has a dense stream and snow station network with over 300 U.S. Geological Survey (USGS) stream stations and around 45 U.S. Department of Agriculture SNOTEL snow stations displayed in Figure 1. For the period of this study, WY 1983, daily observations are available for both stream and snow stations.

Integrated Hydrologic Model

We use the integrated hydrologic model PF-CLM (Maxwell and Miller 2005; Kollet and Maxwell 2006; Maxwell 2013) for all simulations. ParFlow integrates groundwater flow with overland flow and is coupled to a land surface model (CLM) to resolve the energy balance and water balance from the canopy to the ground surface. Together, PF-CLM captures water storage zones and fluxes from the bedrock to the top of the canopy, including overland flow, soil moisture, ET, groundwater flow, and snow processes.

Details about connections between groundwater, surface water and land energy fluxes in ParFlow-CLM have been documented extensively in Maxwell and Miller (2005), Kollet and Maxwell (2006), Kollet and Maxwell (2008), Kollet et al. (2009), Maxwell et al. (2014), Jefferson and Maxwell (2015), Maxwell and Condon (2016), Jefferson et al. (2017). This is a three-dimensional (3D) distributed surface and subsurface water model which solves the saturated and unsaturated Richard's equation together with the kinematic wave equation for surface water transfers over a terrain following grid. This nonlinear system is solved using a globally implicit approach and high parallel efficiency. ParFlow uses a free surface overland flow boundary condition to connect the groundwater and surface water systems. Complete details and testing of this approach are provided by

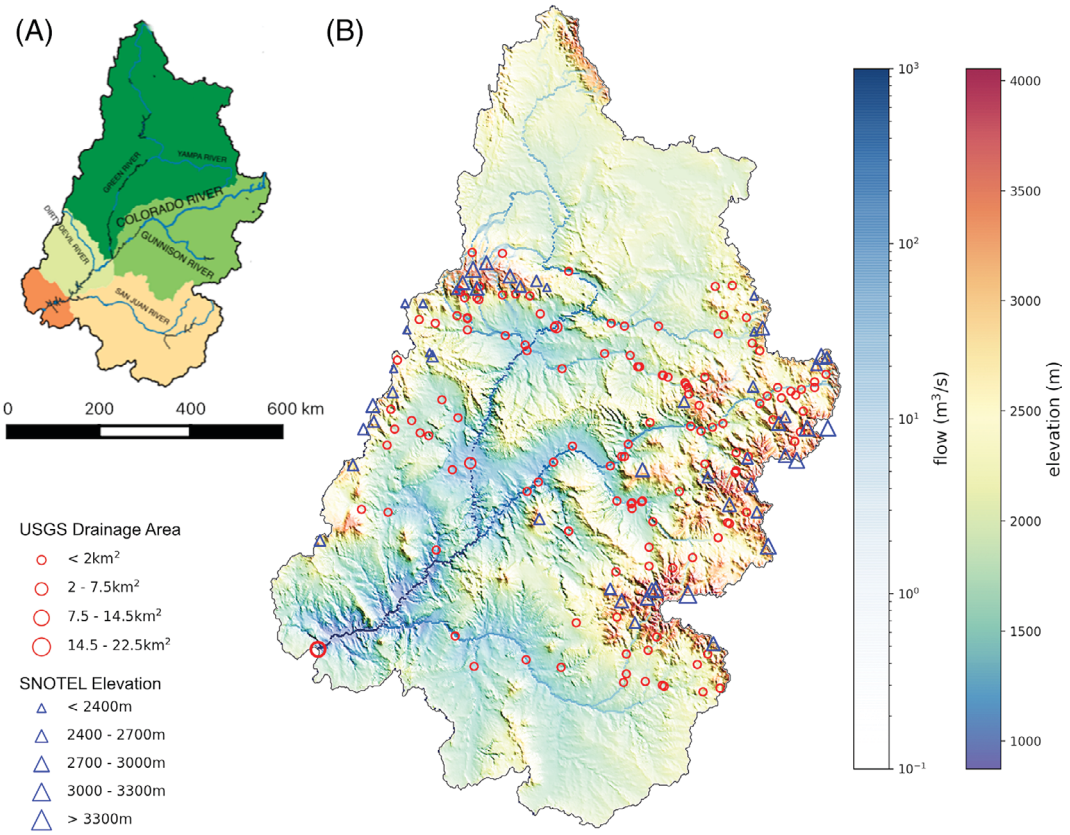

Figure 1. (A) Main rivers and subwatersheds of the Upper Colorado River Basin. (B) The Upper Colorado River Basin and the USGS and SNOTEL stations network. The basin is presented by a stream mask overlain on an elevation map. USGS stations are represented by red circles which their size corresponds to the station's drainage areas. SNOTEL stations are represented by blue triangles which their size corresponds to the station's elevations.

Kollet and Maxwell (2006). Within this framework, continuity of pressure is maintained between subsurface and surface of the domain. This makes it possible to calculate the associated water fluxes at each grid cell using hydrodynamic parameters regardless of the pressure gradient direction (i.e., saturated surface cells can drive infiltration or saturated subsurface pixels exfiltrate water to the surface). With this approach we do not designate stream cells a priori, rather hydrodynamic principles govern where streams/rivers formed via either Hortonian or Dunne runoff (i.e., excess infiltration [Horton 1933] or excess saturation [Dunne 1983]). With this setup, rivers do not incise the subsurface layers. Rather, they form (or disappear) when there is ponding at the surface of the model. Rivers are resolved at the same spatial resolution as the rest of the domain.

Land surface interaction processes are simulated using the CLM which solves ground and vegetation temperature fields with respect of the energy balance at the land surface (Equation 1).

$$R_{net} = H + LE + G \quad (1)$$

where R_{net} is net radiation (W/m^2), H is sensible heat (W/m^2), LE is latent heat (W/m^2) and G is ground heat (W/m^2). The ground heat flux term is solved by the one-dimensional heat conduction equation (Kollet et al. 2009). R_{net} is calculated through a two-sources radiative scheme (soil surface & vegetation temperatures). H and LE are calculated through a resistance scheme including soil, vegetation, and atmospheric resistances (Dai et al. 2003) All of the energy flux terms (H, LE, G) depend directly or indirectly on the water content at or close to the ground surface (Kollet and Maxwell 2008). As such, the available soil moisture calculated by ParFlow is provided to CLM at every timestep to simulate these limitations. ParFlow-CLM couples the subsurface and land-surface system mainly through the general sink/source term and the water availability-controlled energy (Kollet and Maxwell 2008). Please note that plant water availability is controlled by both water at the land surface and soil moisture which depends on lateral groundwater flow and infiltration processes (Maxwell and Condon 2016). For additional technical details on the model please refer to the Supporting Information.

Simulation Domain

Simulation domain for the UCRB is based on the CONUS-ParFlow model (Maxwell et al. 2015) with some additional improvements that are described here. The domain has a spatial resolution of 1 km and five subsurface layers with thicknesses of 0.1, 0.3, 0.6, 1.0,

and 100 m, respectively, from the surface to the bottom. The domain dimensions are 608 km width, 896 km length and 102 m depth, which is equivalent to more than 2.7 million computation cells. We specified no-flow boundary conditions for the bottom and sides of the domain so there are no lateral inflows across the boundaries. The inflows and outflows come from land-surface processes at the top of the domain. Pressures in any cell are calculated at the cell center and can be positive or negative. If the water table falls below the center of the deepest layer this cell would have a negative pressure, but we still have the hydraulic gradients to drive flow.

Subsurface inputs (e.g., soil, geology and bedrock hydraulic characteristics) are selected from the original dataset from Maxwell et al. (2015). Maxwell et al. (2015) assembled the soil dataset from Schaap and Leij (1998) for the top 2 m and from a global permeability map developed by Gleeson et al. (2011) for the deeper subsurface. Then, they extracted hydraulic properties of soil texture information from the soil survey geographic database (SSURGO).

Topographic inputs were developed using a 1-km resolution elevation data upscaled from 30 m National Elevation Dataset (https://ned.usgs.gov) and the stream network raster is developed from the National Hydrography Dataset (NHDPlus). The PriorityFlow toolbox (Condon and Maxwell 2019) was used to process the Digital Elevation Model ensuring a hydrologically consistent drainage network and providing some smoothing along the river network.

Numerical Experiments

Two main numerical experiments are conducted in order to isolate the impact of lateral groundwater flow to surface water in the UCRB: a baseline simulation with no lateral groundwater flow (termed *no lateral* from here on) and a simulation with lateral groundwater flow (termed *lateral*). The *no lateral* simulation is created by: (1) removing topographically driven lateral flow by setting $\sin\theta$ to 0 in Equation S4 and (2) removing the pressure gradient driven lateral flow by decreasing the horizonal hydraulic conductivity in the x- and y-directions ($\boldsymbol{K}_s$ = 1e-5 in x- and y-directions in Equation S4). Both simulations still include groundwater storage, vertical exchanges with the land surface, and lateral surface flow.

Streamflow, lateral groundwater flow and ET are all sensitive to K parameters (Srivastava et al. 2014). Therefore, in addition to the *lateral* and *no lateral* simulations we also evaluate the sensitivity of our findings to the K value used. Foster and Maxwell (2019) scaled K over two orders of magnitude for three respective layers of a ParFlow-CLM model of the East River in Colorado, namely soils, geology and basement layers. They concluded that model outflows are most sensitive to K changes in the geology. Here we repeated our lateral simulations using two uniform K values of $0.1K$ and $10K$ (m/h) comparing to the baseline case of $1K$ in *lateral* configuration setting. While we recognize that perturbing K uniformly is a simplistic sensitivity analysis, our goal with this approach is to understand how important the uncertainty in K was compared to the inclusion and exclusion of lateral groundwater flow.

All simulations are initialized with a completely dry domain and a spinup simulation period is first conducted where we apply constant recharge flux at the land surface until the model reaches steady state. This constant recharge flux was calculated as the long-term average (1950 to 2000) residual of precipitation and evaporation based on datasets developed by Maurer et al. (2002). Once equilibrium is achieved with the long-term average forcing, we then simulate three water years of transient simulation with hourly NLDAS-2 historical data, including incoming short wave and long wave radiations, precipitations, temperature, moisture and pressure of the atmosphere at the surface and wind velocity, to bring the model into a dynamic equilibrium.

We use several metrics to evaluate model performance. Nash-Sutcliffe efficiency (NSE) is a popular metric used to evaluate hydrologic model performance (Nash and Sutcliffe 1970). However, we chose not to use this metric here because it has been shown to be overly sensitive to extreme values (Legates and McCabe Jr. 1999; Krause et al. 2005). NSE can exhibit biases in both time and magnitude (e.g., time-offset or magnitude bias; McCuen et al. 2006). Furthermore, any single metric of model evaluation creates difficultly when diagnosing different sources of bias across a model simulation workflow. Therefore, we use two evaluation metrics to evaluate model performance, namely, Spearman's rank correlation coefficient (Spearman's rho) and total annual flow bias, the combined plotting of these two metrics is known as the Condon diagram (Maxwell and Condon 2016). Spearman's rho is used to evaluate differences in streamflow timing while total bias measures differences in streamflow volumes. Spearman's rho assesses monotonic relationships between two variables without assuming a linear relationship. It is computed as:

$$srho = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} \tag{2}$$

where d_i is the difference in the independent ranking for the simulated and observed values on a given day, and n is the number of values in each time series. Total annual flow is calculated as:

$$bias = \frac{\left|\sum_{i=1}^{n} S_i - \sum_{i=1}^{n} O_i\right|}{\sum_{i=1}^{n} O_i} \tag{3}$$

where S_i and O_i are simulated and observed daily flow on day i and n is the number of days in the WY 1983 (i.e., 365 for this simulation).

Results and Discussion

Our results focus on simulated average annual flows between the *lateral* and *no lateral* simulations to understand spatial differences over the WY1983. We compare

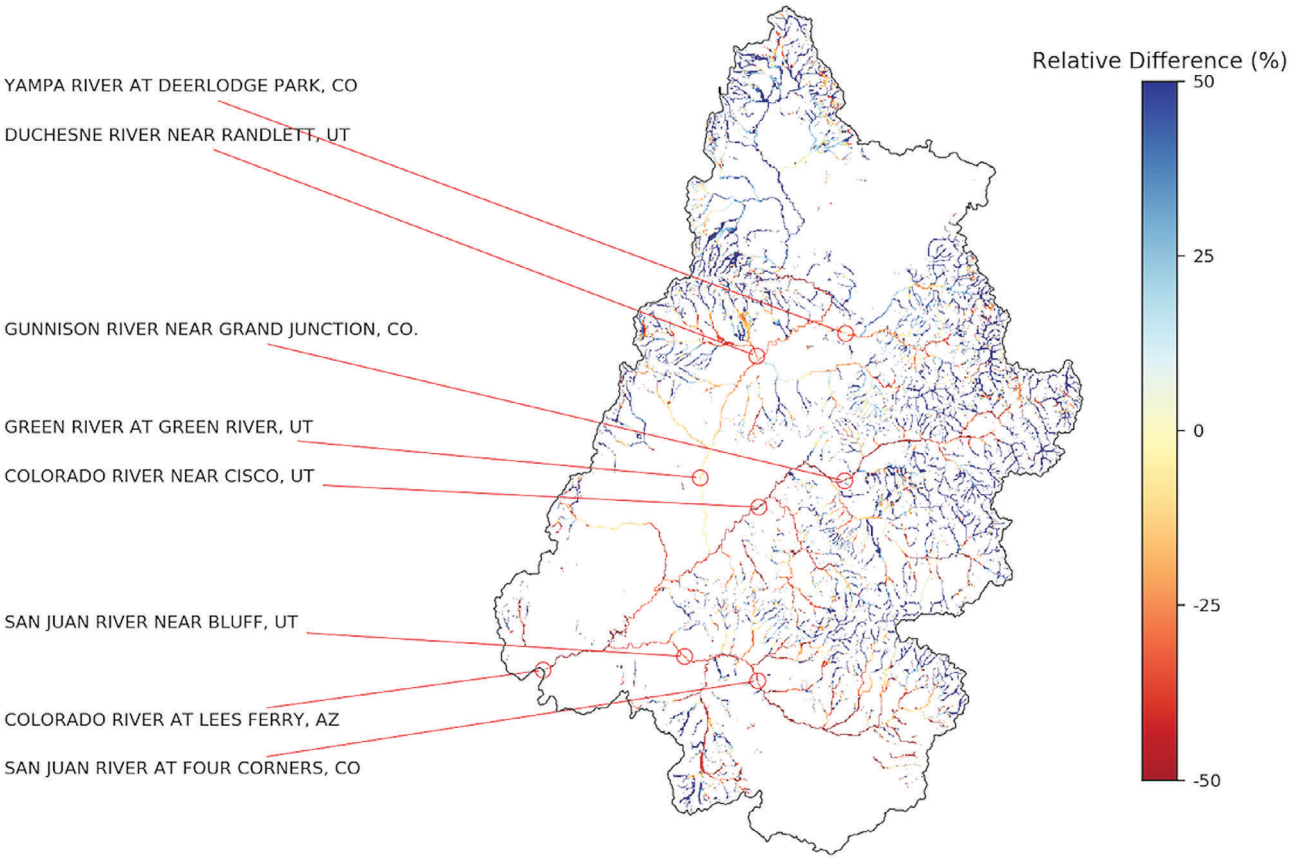

Figure 2. The relative difference between flows from no lateral and lateral simulations. The red circles on the map represent selected stations for time series of flow comparison shown in Figure 3.

these two simulations with observations. Next, we compare model outputs from different hydraulic conductivities to evaluate the sensitivity of lateral groundwater flow to simulations. We then discuss sources of bias in the model outputs. Finally, we evaluate the importance of lateral groundwater flows in peak, total flow and ET by major sub-basins.

Impact of Lateral Groundwater Flow on Streamflow

Figure 2 shows a map of the relative difference in the average annual flows from the *no lateral* and *lateral* simulations. In this figure, we see large differences between simulations in headwater streams. In these headwater streams, flows from the *no lateral* simulation are 20% to 40% greater than flow in the *lateral* simulation. Conversely, in most of the larger rivers, the *lateral* simulations produce 25% to 50% more flow than the *no lateral* simulation. Sensitivity to lateral flow also varies as a function of topographic relief, and most rivers exhibit a point where the flows from the two simulations cross (i.e., the net impact of including lateral groundwater flow swaps between increasing and decreasing flows).

Increased streamflow in the *lateral* case is caused by hydraulic gradients in the subsurface which drive flow from higher elevation recharge locations to areas of groundwater convergence along the stream. However, streamflow increases resulting from lateral convergence can also be counteracted by differences in groundwater configuration between the two cases. The water table tends to be deeper in the higher elevation steep portions of the domain for the *lateral* case relative to the *no lateral* case. These results in higher infiltration and less runoff relative to the *no lateral* case.

Figure 3 plots both the *no lateral* and *lateral* simulated hydrographs and observed flow for eight gauges. We see some trends in this figure. In rivers originating from the Green River plateau, the Duchesne, the Yampa and the Green River, the timing of the *lateral* and *no lateral* results agree.

The effect of lateral groundwater flow is more obvious in rivers originating from the Eastern part of the basin. In the Gunnison River, increases in baseflow and peak flow are 150% and 16%, respectively, from the *no lateral* case to the *lateral* case. In the Colorado River, the corresponding increases are 400% and 38% for the station near Cisco, UT and 430% and 50% for the station at Lee's Ferry, AZ. Streamflow timeseries for two stations in San Juan River show the clearest differences between cases. Here, increases in baseflow and peak flows are around 900% and 57%, respectively, from the *no lateral* simulation to the *lateral* simulation (Figure 3).

Differences in the headwaters of the domain are also translated to downstream streamflow with some timing lags. At the outlet of the UCRB at Lee's Ferry, the peak

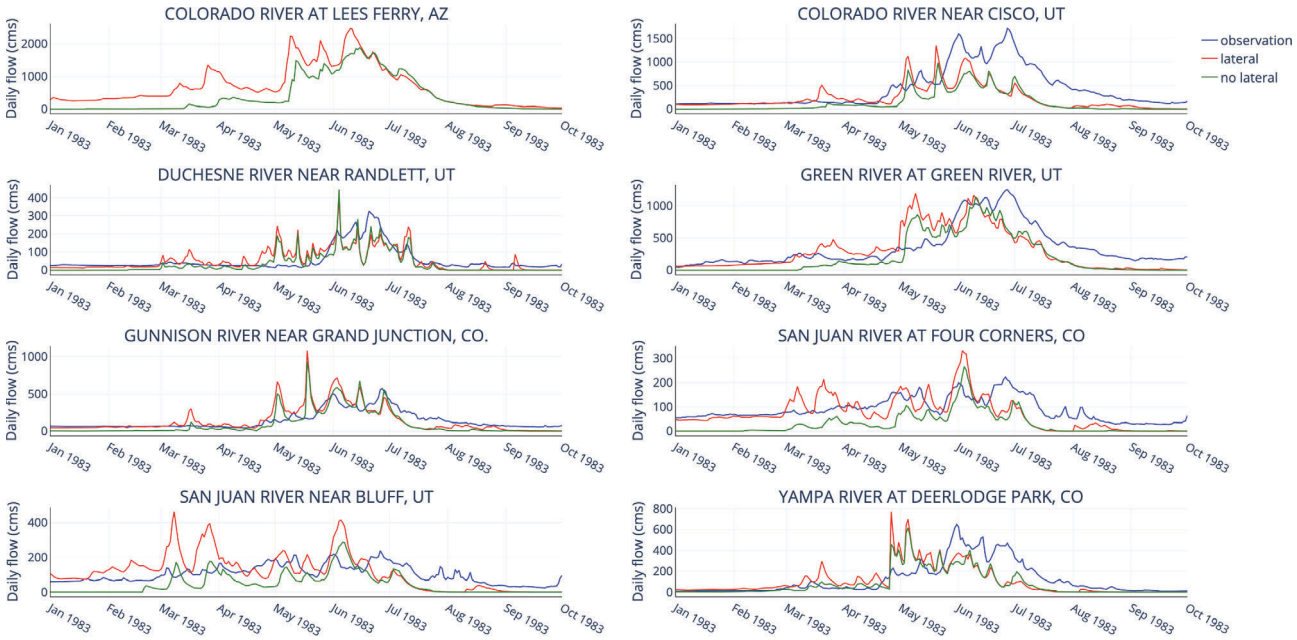

Figure 3. Stream flow time series for the eight representative stations shown in Figure 2. Blue lines are observed daily flow, red and green lines are simulated flows from lateral and no lateral simulation, respectively.

is lagged by 3 days from the *no lateral* case to the *lateral* case. While *lateral* flow peaks at around 2200 m^3/s on May 9, *no lateral* flow peaks at 1480 m^3/s on May 12. The delay continues for the later peaks on May 25 and June 10 for *lateral* flow and on May 28 and June 13 for *no lateral* flow. Indeed, including lateral groundwater flows not only increases base flows but also leads to faster peaks.

In addition to increasing streamflow through exfiltration, the *lateral* simulation also has faster peak times. This is because lateral groundwater flow makes layers close to the surface more saturated than in the *no lateral* case thus resulting in faster flood peaks. However, in complex system like the UCRB, outlet flows are composed from rivers with different characteristics.

At Lee's Ferry flows from the *no lateral* case are lower than from the *lateral* case before June 23 but surpass them after. In the Green River, flows from the two case are quite similar in peak time and magnitude. However, during a recession period starting June 21, the *no lateral* simulation produces more flows than the *lateral* simulation.

Streamflow Comparisons to Observations

All of the simulations presented here are predevelopment simulations, thus, dam and reservoir operations are not incorporated in the modeling process. We therefore focus on USGS stations that are minimally affected by resources management process. However, we still include the Lee's Ferry station in this analysis because it is geographically important as the outlet of our domain. We do not compare to observations for this site as it is situated downstream of Glen Canyon Dam and we are not simulating dam operations. For this station, we only compare simulated results.

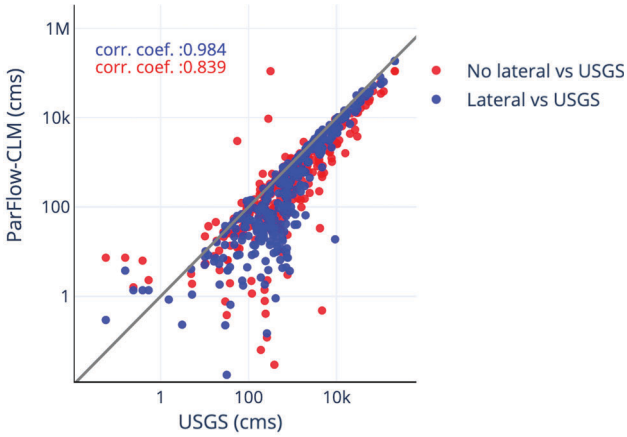

Figure 4. Scatter plots for average annual flows between observations and simulations for all USGS stations in the domain. Blue dots symbolize comparison of flows from lateral simulation with observed ones. Red dots indicate comparison of flows from no lateral simulation with observed flows. Note the log scale used for the axes in this figure.

Figure 4 plots the predicted and observed total flow for all gaging stations in the domain for the *lateral* and *no lateral* cases in log scale. Overall, both simulations demonstrate good model fitness with residual values of 0.839 for the *no lateral* simulation and 0.984 for the *lateral* simulation, respectively. We also see that the *no lateral* flow case produces less streamflow than the *lateral* cases.

The *no lateral* case (red dots) shows much greater underestimation than the *lateral* simulation (blue dots) for stations that have a mean annual flow greater than 50 m^3/s (namely USGS stations along the Colorado River, namely, near Dotsero, CO; Glenwood Spring, CO; near De Beque,

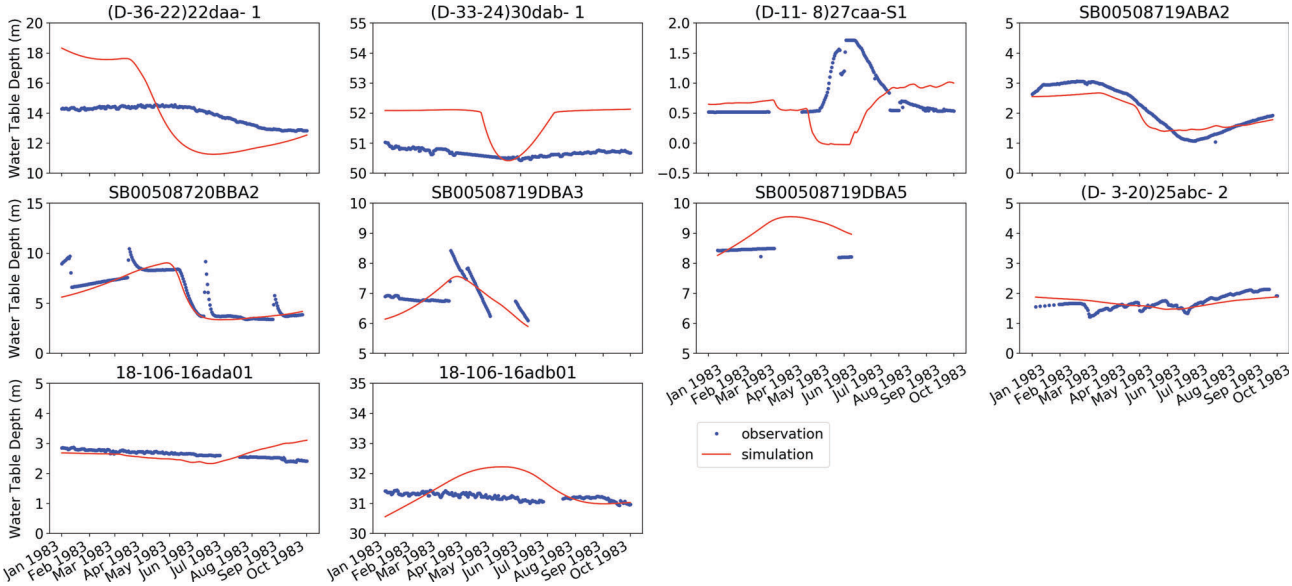

Figure 5. Water table depth time series for the 10 observation wells. Blue dots are observed daily water table depth, red line is simulated water table depth.

CO; near Cameo, CO; near Colorado-Utah state line; near Cisco, UT). This suggests that including lateral flow in the simulation helps to reduce bias between 10% and 50% for stations with mean annual flow greater than 50 m^3/s. It is important to note that this analysis only evaluates total flow, not streamflow timing.

Water table depths from 10 groundwater wells within the domain were compared to the *lateral* model simulations (Figures 5 and 6). These wells all contained transient observations over the simulation period and range in depth from 0.5 to over 50 m. Only 10 wells were available for comparison, but a good agreement is demonstrated between model and observations temporally (Figure 5) and spatially (Figure 6).

We evaluate streamflow performance from lateral simulation with observations for all stations inside the domain using two metrics in tandem: Spearman's rho and total annual flow bias. Follow the Condon-Diagram approach, we scatter-plot the performance results with total annual flow bias and Spearman's rho as horizontal and vertical axes, respectively (Maxwell and Condon 2016). We use the rho value of 0.5 to divide results into good shape (rho greater or equal than 0.5) or bad shape (rho smaller than 0.5). We classify results with bias smaller or equal than 1 as low bias and ones with bias greater than 1 as high bias. Hence, results of this analysis fall into four types: (1) stations with good shape and low bias—green stations; (2) stations with good shape and high bias—blue stations; (3) stations with bad shape and low bias—orange stations; (4) stations with bad shape and high bias—red stations.

Figure 7a shows streamflow observations compared to the *lateral* simulation. Overall, most of the stations are classified as green (50%), the next largest group is low rho and low bias (orange) stations (48%). Figure 7b maps these four performance categories spatially over the domain. The *lateral* simulation performs well for the Green, Gunnison, and San Juan River. For most of the stream origins, the *lateral* simulation reduces the total flows and thus improves correlation with observations from stations in these areas. In addition to the large water projects along the UCRB, many smaller impoundments exist upstream in the system that may also impact the timing of downstream flows. For the representative stations in Figure 2, stations at Green, Yampa, San Juan, and Duchesne Rivers have good shape and low bias. Stations at the Colorado River have poor shape and low bias.

The validation result is consistent with a similar validation of ParFlow outputs over the CONUS domain for the water year 1985 conducted by Maxwell and Condon (2016). Their validation results generally showed agreement between simulated and observed streamflow for the majority of stations (Figure S8). Over the UCRB, Maxwell and Condon (2016) also found a similar distribution at locations with "poor shape, low bias" (Figure S11). They argued that the "poor shape" is due to biases of temperature and precipitation in complex topography regions.

Sensitivity of Simulated Flows to Saturated Hydraulic Conductivity

Figure 8 plots simulated streamflow for the *lateral* (as a baseline), $10K$ and $0.1K$ scenarios. While changing K by two orders of magnitude, simulated flows from two scaled cases are marginally different in magnitude from the original case. Indeed, flows in the two cases behave similarly as observed by Foster and Maxwell (2019). When K was decreased by one order of magnitude, we see lower baseflow and higher peak flow compared to the baseline simulation. When K was increased by an order of magnitude, we see increased baseflow and decreased peak

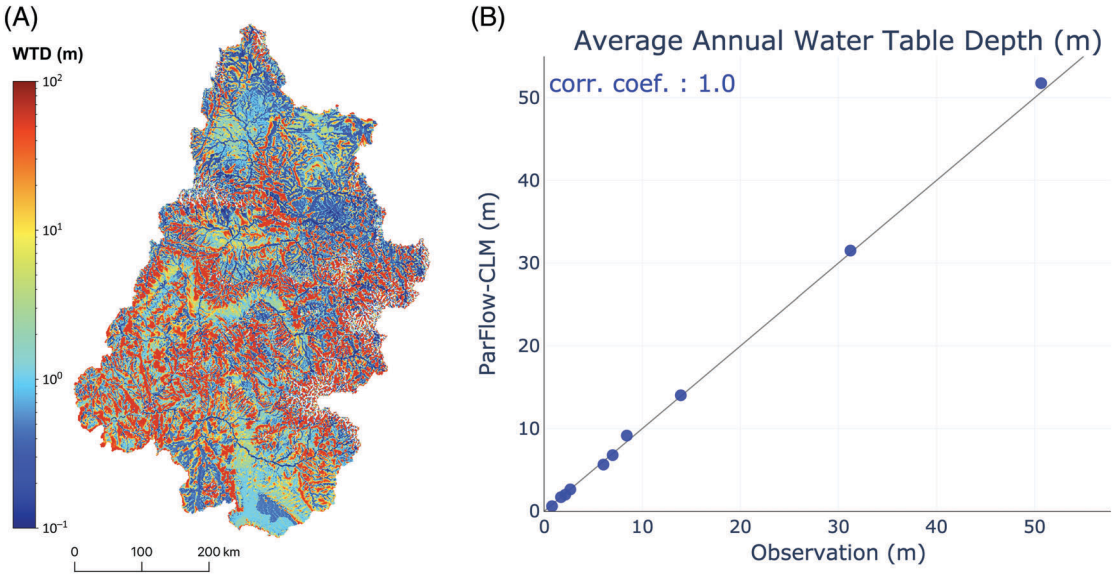

Figure 6. (A) Map of annual average water table depth (m) and (B) scatter plots for average annual water table depth between observations and simulations for all observation wells in the domain.

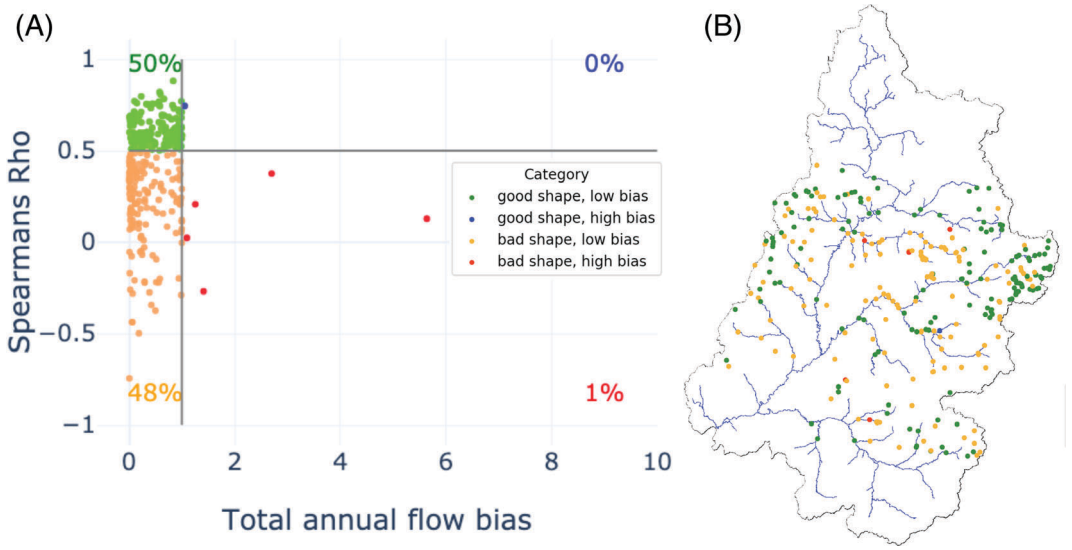

Figure 7. (A) Scatter plot of Spearman's rho and total annual flow bias for every USGS station in the domain. Cutoffs for "good" rho value and "low" bias are indicated with horizontal and vertical lines respectively. Using these thresholds green points are designated as "good shape, low bias," blue are "good shape, high bias," orange are "poor shape, low bias," and red are "poor shape high bias." (B) Map of the corresponding performance categories.

flow. In comparison with the baseline simulation, average changes over the eight stations in peak magnitude are -8% and 3% for $10K$ and $0.1K$ scenarios, respectively. No shift in peak time is observed with changes in K.

Please note that we acknowledge that hydraulic conductivity is an important hydrologic parameter, thus can be a great source of uncertainty. A greater impact by including lateral groundwater flow does not hinder the role of the hydrologic parameters.

Sources of Bias

The performance of large-scale model simulation results may be affected by a number of factors (Maxwell and Condon 2016). Sources of bias include model physics, meteorological forcings, grid resolution and water management operations. Here we discuss some of the major sources of bias and the reasoning behind our model design:

1. *Grid resolution*: While 1 km may be considered coarse resolution in some contexts we chose this resolution here to balance physical complexity with computational demand and available inputs. ParFlow is a computationally efficient model designed to run in parallel, however, solving 3D variability saturated flow in the subsurface is computationally expensive.

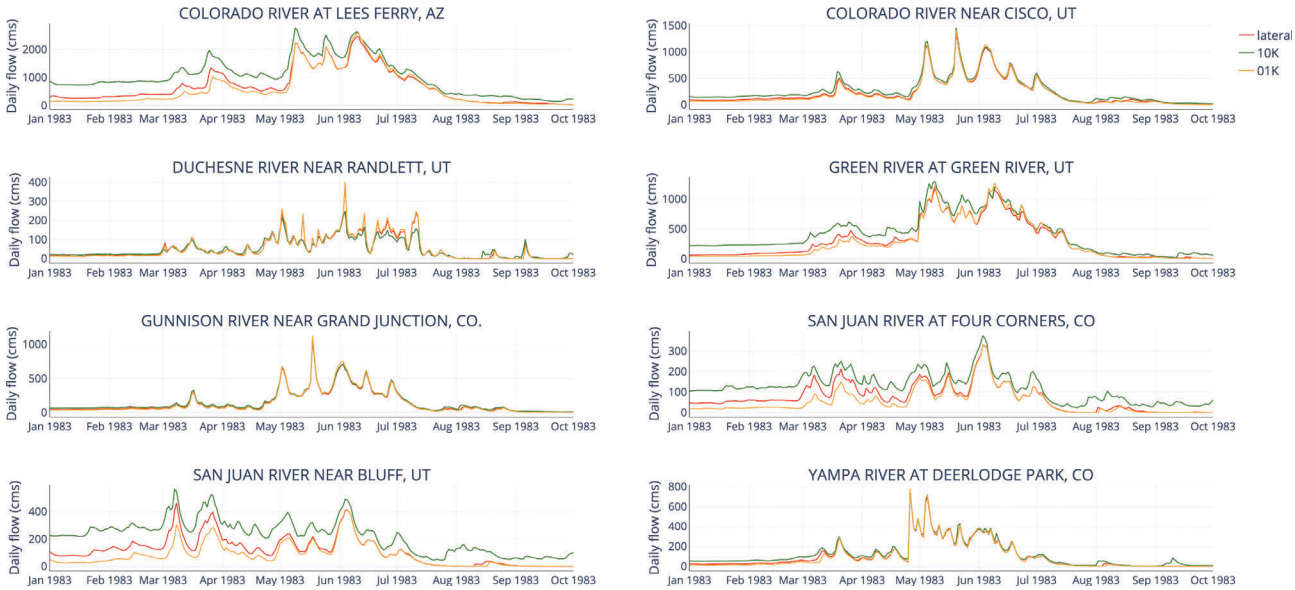

Figure 8. Stream flow time series for the eight representative stations shown in Figure 3. Blue lines are observed daily flow, red lines are simulated flows from lateral simulation, green and orange lines are simulated flows with hydraulic conductivity scaled by 0.1 and 10, respectively.

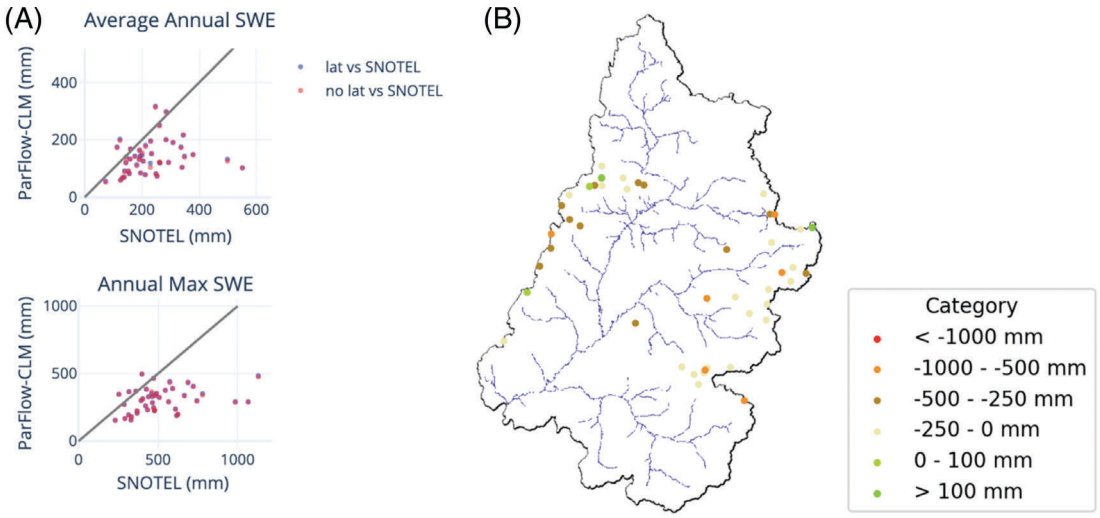

Figure 9. (A) Scatter plots of lateral simulation (blue) and no lateral simulation (red) versus observed (SNOTEL) average and annual max SWE in water year 1983. (B) Map of the bias performance for average annual SWE. Red points/stations have bias smaller than −1000 mm, orange points have bias between −1000 and −500 mm, dark brown points have bias between −500 and −250 mm, light brown points have bias between −250 and 0 mm, light green points have bias between 0 and 100 mm, and green points have bias greater than 100 mm.

Additionally, our spatial resolution is also dictated by available model inputs. High resolution datasets are available for topography and land cover) but geology and forcing data is much more limited. We also acknowledge that appropriate resolutions over which Richards' equation should be applied. Notably, the use of the van Genuchten parameterization requires fine vertical resolution (e.g., Ippisch et al. 2006; Or et al. 2015) which is provided in our simulations. However, much of the lateral flow is saturated, not unsaturated, and this is represented by our approach.

2. *Water management operations*: Our simulations represent predevelopment conditions and therefore do not include surface reservoirs, groundwater pumping, and irrigation. Because we are focusing on peak flows in a record wet year the effect of groundwater pumping and irrigation are less important. However, surface reservoirs would alter peak flow timing and magnitude. We see this bias in our validation results where model performance with respect to both the magnitude and timing of peak flows is better at headwater gauges (less likely to be disturbed) as compared to downstream gauges.

3. *Meteorological forcing data*: NLDAS-2 forcing data applied here is the most comprehensive nationally consistent forcing dataset. However, multiple studies have documented its biases in this dataset (Pan et al. 2003; Sheffield et al. 2003). In order to explore the impact of forcing bias, we compare the simulated Snow water equivalent (SWE) with observations from SNOTEL. Figure 9 shows the comparison between simulated and observed SWE. In this figure we see a systematic underestimation of SWE.
4. As seen from Figure 9a, simulated snowpack for both the *no lateral* and *lateral* simulations is systematically lower than observed. Out of 44 SNOTEL model observation comparisons, only three stations have positive bias; the remaining stations underestimate SWE by 250 to 1000 mm. This is likely due to the coarse NLDAS-2 resolution which has a knows dry warm bias for high elevations. This leads to (1) lower flow volume in Colorado, Green and San Juan Rivers; and (2) earlier snow melt results in earlier simulated flood peaks as is illustrate din the Yampa River at Deerlodge Park, CO and in Green River at Green River streamflow plotted in Figure 3. We acknowledge that there is a scale inconsistency when comparing simulated SWE with 1 km resolution to SNOTEL measurements on a much smaller scale. Nevertheless, SNOTEL is one of the few complete snow observations for the year 1983. Each of the SNOTEL station is mapped to the closest grid cell center of the UCRB domain. The scale inconsistency could be a reason for the bias in validation; however, as can be seen from Figures 1 and 9b, most of the stations that show bias often located at high altitude from 2400 m to over 3300 m. During the record wet year of 1983, snow distribution for 1-km^2 model grids that contain these stations could be considered as uniform in order to compare with simulation results.
5. *Subsurface properties*: Spatially distributed subsurface properties are difficult to obtain. As such, hydraulic conductivity is generally obtained using formal or informal parameter estimation techniques. In the previous section we explore the sensitivity of our results to our choice of hydraulic conductivity values. This demonstrates the uncertainty in our results caused by uncertainty in K as well as the ways in which high or low K may bias our findings.

Differences Between Subwatersheds

The UCRB encompasses many subwatersheds with different relief and other hydrologic characteristics. We explore spatial differences in model performance by considering two different behavior for steeper and flatter portions of the domain. Figure 10 shows relative differences in outputs from the *no lateral* and *lateral* simulations for four subwatersheds delineated in Figure 1. We select in each subwatershed the USGS station, which has the greatest drainage area within the subwatershed as our point of comparison. Simulated flows and ET obtained from these stations' locations are plotted here as the subwatershed value. In the *lateral* simulation, flows are higher across all the subwatershed by 25% to 50%. Differences in peak flow vary more with relief. In flatter subwatersheds (e.g., the Green River), peak flow from the *no lateral* simulation is higher by 18% than peak relative to the *lateral* simulation. In contrast, for regions with steep slopes the peak flows from the *lateral* simulation are considerably higher than the *no lateral* simulation (15% to 30%). However, differences largely diminish at the outlet. The simulated flood peaks between the *no lateral* and *lateral* cases are around 5% at the outlet of the UCRB.

Figure 10d shows the relative difference in distributed average annual ET for the simulations. The *lateral* case has shallower water tables along the river corridor than the *no lateral* case. This shallow groundwater can support ET. In the *no lateral* case, water tables are deeper in these locations because there is no lateral groundwater flow to support convergent zones and as a result ET is lower. Averaging ET differences across subwatersheds the *no lateral* case is 10% to 15% lower than the *lateral* case (Figure 10c). These differences in ET are not as large as the differences in total flow. This suggests a combination of physical routing and soil moisture-ET feedbacks are responsible for the annual flow differences. While average annual flows change from 25% to 50% between two simulations, ET from the *no lateral* simulation is only lower by 8% to 15% than ET from the *lateral* simulation.

Conclusions

Here we study the impact of lateral groundwater flow on simulated peak and annual streamflow in the UCRB using an integrated hydrologic model. For a flood year driven by a very large snowpack, our model simulations suggest that peak flows increase up to 57% when lateral groundwater flow processes are included. Moreover, when compared to a sensitivity analysis where hydraulic conductivity was varied over two orders of magnitude the role of lateral groundwater flow still remains important. Lastly, we see distinct differences in the magnitude and spatial distribution of simulated streamflow with and without lateral groundwater flow between sub-basins of the UCRB.

In flatter rivers such as the Duchesne, the Yampa and the Green, we see the *no lateral* case producing between 10% and 20% less peak flow compared to the *lateral* case. In the steeper river systems, such as the Gunnison, the Colorado and the San Juan, these corresponding differences in peak flows is between 20% and 60%. For the Green River, the differences between *lateral* and *no lateral* cases are smaller than for other sub-basins. This variation between sub-basins highlights the complexity and importance of groundwater lateral flow.

A sensitivity test is carried out to compare the role of lateral groundwater flow changes in hydraulic conductivity. Changes in baseflow and peak flow for both for this sensitivity analysis were small compared to the *lateral* simulation. The differences in the ET between

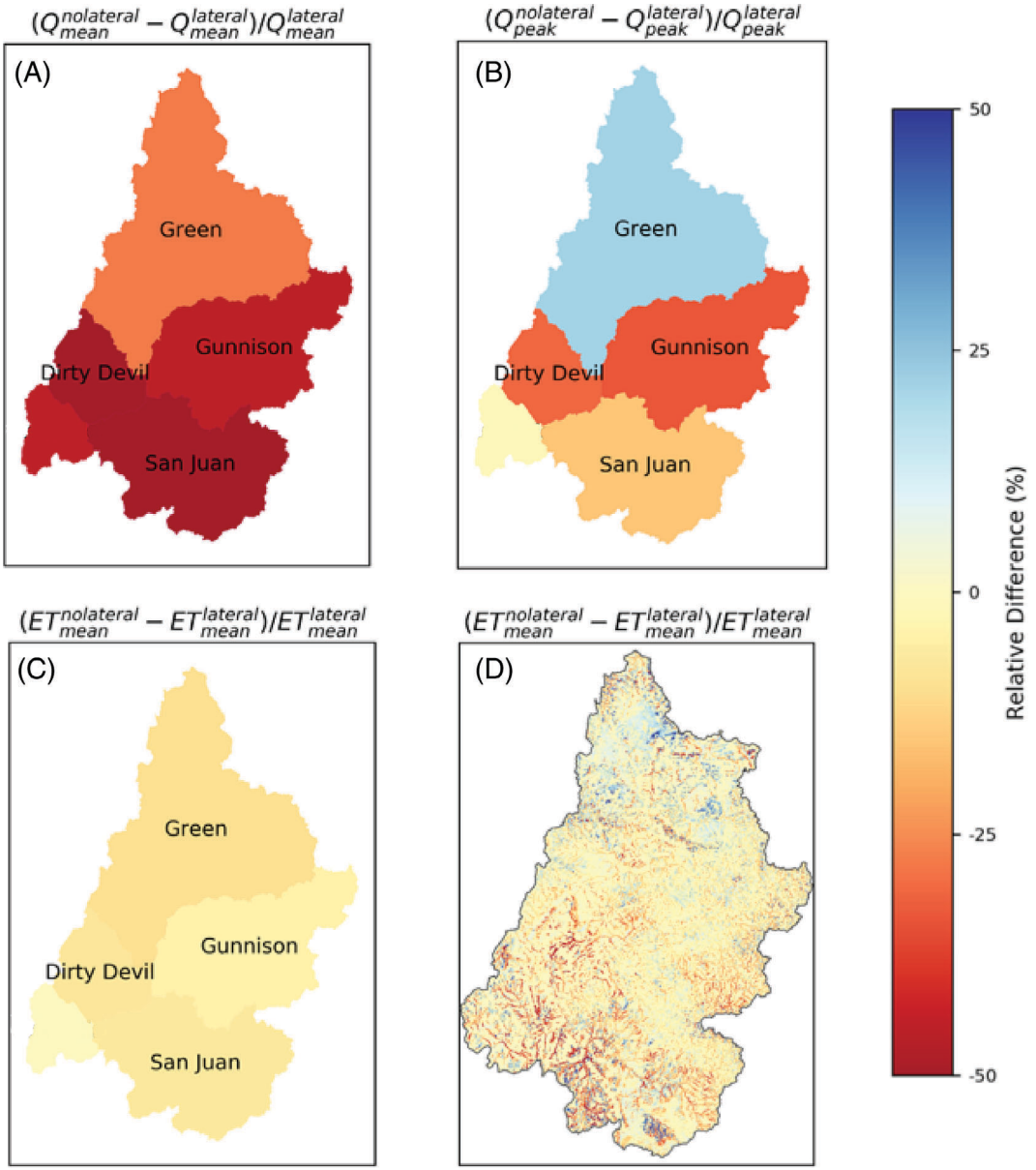

Figure 10. Relative difference between no lateral and lateral simulations for (A) average annual flow, (B) peak flow of each station represents for a subwatershed, (C) average annual ET for each subwatershed, and (D) distributed average annual ET for the domain. The Upper Colorado River Basin is divided into five smaller subwatersheds. In each subwatershed we select the USGS station which has the greatest drainage area and compare simulated flows at these stations' location.

cases averaged across the basin and even across sub-basins cases were small. We see large differences in ET in the river corridors but attribute the difference in streamflow behavior between the *lateral* and *no lateral* cases to physical flow processes, rather than land surface processes.

In large, complex integrated hydrologic simulations there are many sources of bias. These include the meteorological forcing data, model resolution, model equations and formulation, input data and anthropogenic impacts. Precipitation and temperature are likely the most significant among the forcing variables. Snow represents a combined temperature and precipitation signal. As the quantity and timing of snowmelt was particularly important in the water year studied, this forcing bias may have a larger impact on our simulations than other years or other domains. Another source of bias is model resolution, which impacts both the simulated stream network and the magnitude and direction of surface and subsurface lateral flow. While the physics of any model are never perfect, an advantage of the integrated simulation platform used here is the ability to simplify some of the assumptions. In this model (ParFlow) lateral groundwater flow is driven by both pressure and topographic gradients and our work represents an important sensitivity study. While regions of steeper topography influenced our results, other factors were also important. It is important to keep in mind that groundwater flow is driven by gradients in total head

and not just topography. Lateral flow along topographic gradients with shallower water table depths near rivers is an important physical process that increases streamflow. Also, like any model, the input parameters are uncertain and estimated from a range of sources. Saturated hydraulic conductivity is one of the hardest input parameters to quantify and is often estimated in a formal or informal manner. While our simulations may have benefited from a parameter estimation of hydraulic conductivity, the sensitivity analysis does allow us to explore the role this important variable plays on our results. Our simulations are for predevelopment scenarios (i.e., they do not include reservoir operations or pumping). We also attribute some of the differences between simulated and observed streamflow to human activities. Nevertheless, we see a clear improvement in model performance with the inclusion of lateral groundwater flow, for total water (important for water resources) and for peak streamflow (important for flood and reservoir forecasting) and expect this to carry over to integrated simulations that include water management.

As the Colorado River is heavily regulated, it is important to accurately simulate reservoir inflows for exceptional years (drought and flood). Our simulation results suggest that an integrated model of this system may help inform dam operators under conditions such as the winter of 1983 when the Colorado River peaked twice during snowmelt resulting in critical conditions at the Glen Canyon Dam. As we may see more extreme precipitation events combined with more rapid snowmelt under a changing climate (Musselman et al. 2017) these types of integrated model simulations may prove to be a useful tool for water management on the UCRB and including lateral groundwater flow may aid in improving the fidelity of these model projections. In this work and with this model we can better quantify the importance of lateral groundwater flow to surface water.

Acknowledgments

This work was supported by U.S. National Science CyberInfrastructure project, HydroFrame (NSF-OAC 1835903) and the IDEAS-Watersheds (DOE Office of Science) project. All output data and model inputs will be made available through the HydroFrame project with data hosted via CyVerse. ParFlow is an open source modeling platform and the source code used in these simulations and associated documentation may be obtained from https://github.com/parflow/parflow.

Authors' Note

The authors do not have any conflicts of interest or financial disclosures to report.

Supporting Information

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

Appendix S1. Supporting Information.

References

Alexander, R.B., E.W. Boyer, R.A. Smith, G.E. Schwarz, and R.B. Moore. 2007. The role of headwater streams in downstream water quality. *Journal of the American Water Resources Association* 43, no. 1: 41–59. https://doi.org/10.1111/j.1752-1688.2007.00005.x

Bales, R.C., N.P. Molotch, T.H. Painter, M.D. Dettinger, R. Rice, and J. Dozier. 2006. Mountain hydrology of the western United States. *Water Resources Research* 42, no. 8. https://doi.org/10.1029/2005WR004387

Barnett, T.P., J.C. Adam, and D.P. Lettenmaier. 2005. Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature* 438, no. 7066: 303–309. http://doi.org/10.1038/nature04141

Brouyère, S., G. Carabin, and A. Dassargues. 2004. Climate change impacts on groundwater resources: Modelled deficits in a chalky aquifer, Geer basin, Belgium. *Hydrogeology Journal* 12, no. 2: 123–134. https://doi.org/10.1007/s10040-003-0293-1

Buto, S.G., L.E. Spangler, A.L. Flint, and L.E. Flint. 2017. Catchment-flowline network and selected model inputs for an enhanced and updated spatially referenced statistical assessment of dissolved-solids load sources and transport in streams of the Upper Colorado River Basin. *U.S. Geological Survey data release*, https://doi.org/10.5066/F76T0JT4.

Chang, A.T.C., J.L. Foster, P. Gloersen, W.J. Campbell, and E.G. Josberger. 1987. Estimating snowpack parameters in the Colorado River Basin. *International Association of Hydrological Sciences Publication* 166: 343–352.

Christensen, N.S., and D.P. Lettenmaier. 2007. A multimodel ensemble approach to assessment of climate change impacts on the hydrology and water resources of the Colorado River Basin. *Hydrology and Earth System Sciences* 11, no. 4: 1417–1434. https://doi.org/10.5194/hess-11-1417-2007

Christensen, N.S., A.W. Wood, N. Voisin, D.P. Lettenmaier, and R.N. Palmer. 2004. The effects of climate change on the hydrology and water resources of the Colorado River Basin. *Climatic Change* 62, no. 1–3: 337–363. https://doi.org/10.1023/B:CLIM.0000013684.13621.1f

Condon, L.E., and R.M. Maxwell. 2019. Modified priority flood and global slope enforcement algorithm for topographic processing in physically based hydrologic modeling applications. *Computers and Geosciences* 126: 73–83. https://doi.org/10.1016/j.cageo.2019.01.020

Condon, L.E., A.L. Atchley, and R.M. Maxwell. 2020. Evapotranspiration depletes groundwater under warming over the contiguous United States. *Nature Communications* 11, no. 1. https://doi.org/10.1038/s41467-020-14688-0

Dai, Y., X. Zeng, R.E. Dickinson, I. Baker, G.B. Bonan, M.G. Bosilovich, A.S. Denning, P.A. Dirmeyer, P.R. Houser, G. Niu, K.W. Oleson, C.A. Schlosser, and Z.L. Yang. 2003. The common land model. *Bulletin of the American Meteorological Society* 84, no. 8: 1013–1023.

Decharme, B., R. Alkama, H. Douville, M. Becker, and A. Cazenave. 2010. Global evaluation of the ISBA-TRIP continental hydrological system. Part II: Uncertainties in river routing simulation related to flow velocity and groundwater storage. *Journal of Hydrometeorology* 11, no. 3: 601–617. https://doi.org/10.1175/2010JHM1212.1

Dunne, T. 1983. Relation of field studies and modeling in the prediction of storm runoff. *Journal of Hydrology* 65, no. 1–3: 25–48. https://doi.org/10.1016/0022-1694(83)90209-3

Environmental Protection Agency. 2015. Importance and benefits of primary headwater streams. Epa.ohio.gov. https://www.epa.ohio.gov/portals/35/wqs/headwaters/HWH_import.pdf (accessed August 1, 2019).

Fang, X., and J.W. Pomeroy. 2016. Impact of antecedent conditions on simulations of a flood in a mountain headwater basin. *Hydrological Processes* 30, no. 16: 2754–2772. https://doi.org/10.1002/hyp.10910

Flessa, K.W., E.P. Glenn, O. Hinojosa-Huerta, C.A. de la Parra-Rentería, J. Ramírez-Hernández, J.C. Schmidt, and F.A. Zamora-Arroyo. 2013. Flooding the Colorado River Delta: A landscape-scale experiment. *Eos, Transactions American Geophysical Union* 94, no. 50: 485–486. https://doi.org/10.1002/2013EO500001

Foster, L.M., and R.M. Maxwell. 2019. Sensitivity analyses of hydraulic conductivity and Manning's *n* parameters lead to new method to scale effective hydraulic conductivity across model resolutions. *Hydrological Processes* 33, no. 3: 332–349. https://doi.org/10.1002/hyp.13327

Franz, K.J., H.C. Hartmann, S. Sorooshian, and R. Bales. 2003. Verification of national weather service ensemble streamflow predictions for water supply forecasting in the Colorado River Basin. *Journal of Hydrometeorology* 4, no. 6: 1105–1118. https://doi.org/10.1175/1525-7541(2003)004%3C1105:VONWSE%3E2.0.CO;2

Gleeson, T., L. Smith, N. Moosdorf, J. Hartmann, H.H. Dürr, A.H. Manning, L.P.H. van Beek, and A.M. Jellinek. 2011. Mapping permeability over the surface of the earth. *Geophysical Research Letters* 38, no. 2. https://doi.org/10.1029/2010GL045565

Horton, R.E. 1933. The Rôle of infiltration in the hydrologic cycle. *Transactions, American Geophysical Union* 14, no. 1: 446. https://doi.org/10.1029/tr014i001p00446

Ippisch, O., H.-J. Vogel, and P. Bastian. 2006. Validity limits for the van Genuchten–Mualem model and implications for parameter estimation and numerical simulation. *Advances in Water Resources* 29, no. 12: 1780–1789. https://doi.org/10.1016/j.advwatres.2005.12.011

Jefferson, J.L., and R.M. Maxwell. 2015. Evaluation of simple to complex parameterizations of bare ground evaporation. *Journal of Advances in Modeling Earth Systems* 7, no. 3: 1075–1092. https://doi.org/10.1002/2014ms000398

Jefferson, J.L., R.M. Maxwell, and P.G. Constantine. 2017. Exploring the Sensitivity of Photosynthesis and Stomatal Resistance Parameters in a Land Surface Model. *Journal of Hydrometeorology* 18, no. 3: 897–915. https://doi.org/10.1175/jhm-d-16-0053.1

Kollet, S.J., I. Cvijanovic, D. Schüttemeyer, R.M. Maxwell, A.F. Moene, and P. Bayer. 2009. The influence of rain sensible heat and subsurface energy transport on the energy balance at the land surface. *Vadose Zone Journal* 8, no. 4: 846–857. https://doi.org/10.2136/vzj2009.0005

Kollet, S.J., and R.M. Maxwell. 2008. Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model. *Water Resources Research* 44, no. 2: W02402. https://doi.org/10.1029/2007WR006004

Kollet, S.J., and R.M. Maxwell. 2006. Integrated surface–groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. *Advances in Water Resources* 29, no. 7: 945–958. https://doi.org/10.1016/j.advwatres.2005.08.006

Kollet, S., M. Sulis, R.M. Maxwell, C. Paniconi, M. Putti, G. Bertoldi, E.T. Coon, C. Cordano, S. Endrizzi, E. Kikinzon, E. Mouche, C. Mügler, Y.-J. Park, J.C. Refsgaard, S. Stisen, and E. Sudicky. 2017. The integrated hydrologic model intercomparison project, IH-MIP2: A second set of benchmark results to diagnose integrated hydrology and-feedbacks. *Water Resources Research* 53, no. 1: 867–890. https://doi.org/10.1002/2016wr019191

Kopytkovskiy, M., M. Geza, et al. 2015. Climate-change impacts on water resources and hydropower potential in the Upper Colorado River Basin. *Journal of Hydrology: Regional Studies* 3: 473–493.

Krause, P., D.P. Boyle, and F. Base. 2005. Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences* 5: 89–97. SRef-ID: 1680-7359/adgeo/2005-5-89.

Legates, D.R., and G.J. McCabe. 1999. Evaluating the use of 'goodness-of-fit' Measures in hydrologic and hydroclimatic model validation. *Water Resources Research* 35, no. 1: 233–241. https://doi.org/10.1029/1998wr900018

Maurer, E.P., A.W. Wood, J.C. Adam, D.P. Lettenmaier, and B. Nijssen. 2002. Long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. *Journal of Climate* 15: 3237–3251. https://doi.org/10.1175/1520-0442(2002)015%3C3237:ALTHBD%3E2.0.CO;2

Maxwell, R.M., and L.E. Condon. 2016. Connections between groundwater flow and transpiration partitioning. *Science* 353, no. 6297: 377–380. https://doi.org/10.1126/science.aaf7891

Maxwell, R.M., L.E. Condon, and S.J. Kollet. 2015. A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3. *Geoscientific Model Development* 8, no. 3: 923–937. https://doi.org/10.5194/gmd-8-923-2015

Maxwell, R.M. 2013. A terrain-following grid transform and preconditioner for parallel, large-scale, integrated hydrologic modeling. *Advances in Water Resources* 53: 109–117. https://doi.org/10.1016/j.advwatres.2012.10.001

Maxwell, R.M., and N.L. Miller. 2005. Development of a coupled land surface and groundwater model. *Journal of Hydrometeorology* 6, no. 2: 33–47. https://doi.org/10.1175/JHM422.1

Maxwell, R.M., M. Putti, S. Meyerhoff, J.-O. Delfs, I.M. Ferguson, V. Ivanov, J. Kim, O. Kolditz, S.J. Kollet, M. Kumar, S. Lopez, J. Niu, C. Paniconi, Y.-J. Park, M.S. Phanikumar, C. Shen, E.A. Sudicky, and M. Sulis. 2014. Surface-subsurface model intercomparison: A first setofbenchmark results to diagnose integrated hydrology and feedbacks. *WaterResources Research* 50, no. 2: 1531–1549. https://doi.org/10.1002/2013wr013725

McCuen, R.H., Z. Knight, and A.G. Cutter. 2006. Evaluation of the Nash-Sutcliffe efficiency index. *Journal of Hydrologic Engineering* 11, no. 6: 597–602. https://doi.org/10.1061/(ASCE)1084-0699(2006)11:6(597)

Miller, M.P., S.G. Buto, D.D. Susong, and C.A. Rumsey. 2016. The importance of base flow in sustaining surface water flow in the Upper Colorado River Basin. *Water Resources Research* 52, no. 5: 3547–3562. https://doi.org/10.1002/2015WR017963

Musselman, K.N., M.P. Clark, C. Liu, K. Ikeda, and R. Rasmussen. 2017. Slower snowmelt in a warmer world. *Nature Climate Change* 7, no. 3: 214–219. https://doi.org/10.1038/nclimate3225

Nash, J.E., and J.V. Sutcliffe. 1970. River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology* 10, no. 3: 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Nash, L.L., and P.H. Gleick. 1991. The sensitivity of streamflow in the Colorado Basin to climatic changes. *Journal of Hydrology* 125, no. 3–4: 221–241. https://doi.org/10.1016/0022-1694(91)90030-L

Or, D., P. Lehmann, and S. Assouline. 2015. Natural length scales define the range of applicability of the Richards equation for capillary flows. *Water Resources Research* 51: 7130–7144. https://doi.org/10.1002/2015WR017034

Painter, T.H., J.S. Deems, J. Belnap, A.F. Hamlet, C.C. Landry, and B. Udall. 2010. Response of Colorado River runoff to dust radiative forcing in snow. *Proceedings of the National*

Academy of Sciences of the United States of America 107, no. 40: 17125–17130. https://doi.org/10.1073/pnas.0913139107

Pan, M., J. Sheffield, E.F. Wood, K.E. Mitchell, P.R. Houser, J.C. Schaake, A. Robock, D. Lohmann, B. Cosgrove, Q.Y. Duan, L. Luo, R.W. Higgins, R.T. Pinker, and J.D. Tarpley. 2003. Snow process modeling in the north American land data assimilation system (NLDAS): 2. Evaluation of model simulated snow water equivalent. *Journal of Geophysical Research: Atmospheres* 108, no. D22: 8850. https://doi.org/10.1029/2003JD003994

Schaap, M.G., and F.J. Leij. 1998. Database related accuracy and uncertainty of pedotransfer functions. *Soil Science* 163: 765–779. https://doi.org/10.1097/00010694-199810000-0000

Sear, D.A., P.D. Armitage, and F.H. Dawson. 1999. Groundwater dominated rivers. *Hydrological Processes* 13: 3. https://doi.org/10.1002/(SICI)1099-1085(19990228)13:3<255::AID-HYP737>3.0.CO;2-Y

Sheffield, J., M. Pan, E.F. Wood, K.E. Mitchell, P.R. Houser, J.C. Schaake, A. Robock, D. Lohmann, B. Cosgrove, Q.Y. Duan, L.F. Luo, R.W. Higgins, R.T. Pinker, J.D. Tarpley, and B.H. Ramsay. 2003. Snow process modeling in the north American land data assimilation system (NLDAS): 1. Evaluation of model-simulated snow cover extent. *Journal of Geophysical Research: Atmospheres* 108, no. D22: 8849. https://doi.org/10.1029/2002JD003274

Srivastava, V., W. Graham, R. Muñoz-Carpena, and R.M. Maxwell. 2014. Insights on geologic and vegetative controls over hydrologic behavior of a large complex basin—Global sensitivity analysis of an integrated parallel hydrologic model. *Journal of Hydrology* 519, no. PB: 2238–2257. https://doi.org/10.1016/j.jhydrol.2014.10.020

Tillman, F.D. 2015. Documentation of input datasets for the soil-water balance groundwater recharge model of the Upper Colorado River Basin. U.S. Geological Survey Open-File Report 2015-1160, 17 p. https://doi.org/10.3133/ofr20151160.

Vandivere, W.B., and P. Vorster. 1984. Hydrology analysis of the Colorado River floods of 1983. *GeoJournal* 9, no. 4: 343–350. https://doi.org/10.1007/BF00697962

Woodhouse, C.A., and G.T. Pederson. 2018. Investigating runoff efficiency in upper Colorado River streamflow over past centuries. *Water Resources Research* 54, no. 1: 286–300. https://doi.org/10.1002/2017WR021663