Received: 30 September 2021 | Revised: 25 February 2022 | Accepted: 27 February 2022
DOI: 10.1002/hyp.14536

TOWARDS MORE CREDIBLE MODELS IN CATCHMENT HYDROLOGY TO ENHANCE HYDROLOGICAL PROCESS UNDERSTANDING

WILEY

Multicriteria analysis on rock moisture and streamflow in a rainfall-runoff model improves accuracy of model results

Peter T. La Follette[1] | W. Jesse Hahm[2] | Daniella M. Rempe[3] | William E. Dietrich[4] | Claudia C. Brauer[5] | Albrecht H. Weerts[5,6] | David N. Dralle[7]

[1]Lynker Technologies, Boulder, Colorado, USA
[2]Department of Geography, Simon Fraser University, Burnaby, British Columbia, Canada
[3]Department of Geological Science, Jackson School of Geosciences, University of Texas, Austin, Texas, USA
[4]Department of Earth and Planetary Science, University of California, Berkeley, California, USA
[5]Hydrology and Quantitative Water Management Group, Wageningen University, Wageningen, The Netherlands
[6]Operational Water Management, Dept of Inland Water Systems, Deltares, Delft, The Netherlands
[7]Pacific Southwest Research Station, United States Forest Service, Albany, California, USA

Correspondence
Peter T. La Follette, Lynker Technologies, Boulder, CO, USA.
Email: plafollette@lynker.com

Abstract

Although shallow (< 1.5 m) soil water storage has been extensively studied, the significance of deeper unsaturated zone water storage to flow generation is poorly documented. However, a limited but growing body of empirical work shows that the weathered bedrock vadose zone, not soil, stores the majority of plant available water in many seasonally dry and semi-arid landscapes. Moreover, this storage dynamic mediates recharge to hillslope groundwater systems that generate stream discharge and support ecologically significant baseflows. Explicit representations of bedrock vadose zone processes are rarely incorporated into runoff models, due in part to a paucity of observations that can constrain simulations. Here, we develop a simple representation of the weathered bedrock vadose zone that is guided by in situ field observations. We incorporate this representation into a rainfall-runoff model, and calibrate it on streamflow alone, on rock moisture (i.e., weathered bedrock vadose zone moisture) alone, and on both using the concept of Pareto optimality. We find that the model is capable of accurately simultaneously simulating dynamics in rock moisture and streamflow, in terms of Kling-Gupta Efficiency, when using Pareto optimal parameter sets. Calibration on streamflow alone, however, is insufficient to accurately simulate rock moisture dynamics. We further find that the posterior distributions of some model parameters are sensitive to choice of calibration scenario. The posterior distribution of high-performing model parameters resulting from the streamflow only calibration scenario include physically unrealistic values that are not yielded by the rock moisture only or Pareto calibration strategies. These results suggest that the accuracy of some model results can be increased and parameter uncertainty decreased via incorporation of rock moisture data in calibration, without sacrificing streamflow simulation quality. Emerging recognition of the global significance of weathered bedrock water storage in seasonally dry and semi-arid regions motivates more observations of weathered bedrock moisture and integration of this variable into earth system models.

KEYWORDS
conceptual hydrologic model, multiple criteria analysis, parameter uncertainty, rock moisture, vadose zone

1 INTRODUCTION

Water storage is a key variable that governs catchment function (e.g., Kirchner, 2009; McDonnell et al., 2018). For example, storage in the vadose zone typically determines how much water is available for plants and how frequently groundwater is recharged (Botter et al., 2007; Porporato et al., 2004; Rempe & Dietrich, 2018; Salve et al., 2012). Release from groundwater systems can drive floods and sustain ecologically significant baseflow (e.g., Zipper et al., 2019). Yet, compared to relatively common measurements of discharge, direct measurements of catchment storage, especially of the vadose zone moisture below the soil mantle, at spatial scales relevant to hydrological modelling and management are logistically challenging to obtain and are therefore relatively rare.

Without constraints on model parameter ranges used in calibration, models with various representations and parameterizations of the hillslope "black box" (including both its runoff generation mechanisms and storage state) can nonetheless produce adequate predictions of streamflow (Bouaziz et al., 2020; Hartmann et al., 2013; Koch et al., 2016; Koppa et al., 2019; Rakovec, Kumar, Mai, et al., 2016, as judged through error metrics such as the Nash-Sutcliffe efficiency (NSE; Nash and Sutcliffe (1970)) and Kling-Gupta efficiency (KGE; Gupta et al. (2009)). However, other components of the hydrologic system (such as state variables or internal fluxes) may not be sufficiently resolved if the model is constrained on streamflow observations alone. Bouaziz et al. (2020) simulated streamflow in the Meuse basin with 12 hydrologic models with similar performances in simulating streamflow; these models yielded very different internal states, which cannot all simultaneously be accurate. Hartmann et al. (2013) found that there is not enough information in streamflow to identify relevant processes via modelling efforts in a karst system. Rakovec, Kumar, Mai, et al. (2016) found 'that constraining the model against streamflow only may be necessary but not sufficient to warrant the model fidelity for other complementary variables'. This so-called equifinality problem is undesirable when models are used to make predictions under scenarios in which baseline drivers have shifted (non-stationary behaviour, due to, for example, climate change or land use, as detailed by Bouaziz et al. (2021)), and when accurate inferences of internal parameters or state variables are important (Anderton et al., 2002; Bouaziz et al., 2020; Heppner et al., 2007; Hrachowitz & Clark, 2017). Multiple criteria analysis, or MCA, has emerged as a useful tool which can help to reduce parameter uncertainty in hydrologic models and potentially improve accuracy of multiple model results. While MCA has been employed on multiple objective functions of streamflow (Gupta et al., 1998; Pokhrel et al., 2012; Rosolem et al., 2012), recent developments in remote sensing and the observation of hydrologically relevant fluxes and state variables (aside from streamflow) has enabled MCA on novel components of the hydrological system (Koppa et al., 2019; Lopez et al., 2017; Rakovec, Kumar, Attinger, et al., 2016; Rakovec, Kumar, Mai, et al., 2016; Sadeghi et al., 2020; Wanders et al., 2014).

In hillslopes, water storage can occur in the mobile, typically shallow soil, saprolite (which can be defined as a soil-like material that nevertheless remains in situ and retains relict bedrock structures), and within the fractures and pores of weathered bedrock, which has undergone chemical and physical alteration relative to the fresh material at depth 2 (e.g., Graham et al., 2010; Klos et al., 2018). The underlying fresh, unweathered bedrock is typically not considered as a major contributor to streamflow or transpiration (e.g., Rempe & Dietrich, 2014; Riebe et al., 2017). Hillslope scale field studies demonstrate that weathered bedrock commonly routes runoff (e.g., Asano et al., 2002; Blumstock et al., 2015; Frisbee et al., 2011; Hale et al., 2016; Haria & Shand, 2004; Katsuyama et al., 2010; Montgomery et al., 1997; Rademacher et al., 2005). However, water storage within weathered bedrock is largely understood through inferences made from streamflow chemistry (e.g., Soulsby et al., 2007) or soil and groundwater observations (e.g., Appels et al., 2015; Freer et al., 2002; Gabrielli et al., 2012; Safeeq et al., 2021; Smith et al., 2011) and few studies quantitatively observe water storage within weathered bedrock (e.g., Haria & Shand, 2004; Ireson et al., 2006; Jardine et al., 1999; Salve et al., 2012). In addition to routing runoff, the weathered bedrock is thought to commonly serve as an important plant-available water storage reservoir that mediates recharge and runoff (Hahm, Rempe, et al. 2019; Lovill et al., 2018; Salve et al., 2012), and is therefore a key control on catchment water balance dynamics and precipitation partitioning. The role of the bedrock vadose zone is likely amplified in locations where long periods without precipitation tend to increase vegetation reliance on deep water sources (Dawson et al., 2020; Goulden & Bales, 2019; Rose et al., 2003; Schwinning, 2010).

Using methods developed by Kirchner (2009) and Sayama et al. (2011), Dralle et al. (2018) recently observed that seasonally dynamic storage (i.e., seasonal gains and losses in total catchment storage) at a seasonally dry (Mediterranean-type) catchment named Elder Creek (in the northern California Coast Ranges) could be conceptually sub-divided between two reservoirs: one reservoir that directly drove hydraulic gradients leading to streamflow generation (termed 'direct storage'), and a second reservoir that was seasonally dynamic but hydraulically disconnected from the stream (termed 'indirect storage'). Dralle et al. (2018) showed that the latter matched previously observed storage magnitudes (and temporal dynamics of those volumes) in the unsaturated zone, which can be further divided between storage in unsaturated soils ('soil moisture') and unsaturated weathered bedrock ('rock moisture', which we spatially refer to as the 'bedrock vadose zone') (Rempe & Dietrich, 2018; Salve et al., 2012). Surprisingly, the maximum observed dynamic storage in the indirect storage reservoir (approximately 400 mm) far exceeded the volumes in direct storage (approximately 100 mm), which was interpreted to represent the deeper fractured-rock hillslope aquifer system that feeds the stream through a network of seeps and springs (Lovill et al., 2018). Furthermore, of the two components of indirect storage, the seasonal rock moisture storage was found to be 4.7 times larger than the seasonal moisture storage as soil moisture (Rempe & Dietrich, 2018).

The observation that indirect storage can be the largest part of seasonally dynamic storage—and that rock moisture can be the largest component of indirect storage—underscores how rock moisture can play a central role in catchment hydrologic function. Consistent with this role, rock moisture at the site has been argued to sustain dry season forest transpiration (Hahm, Rempe, et al., 2019; Oshun et al., 2012; Rempe & Dietrich, 2018), regulate solute export (Kim et al., 2017), be the locus of root respiration Tune et al. (2020), determine ecosystem sensitivity to drought (Hahm, Dralle, et al., 2019), and mediate recharge to the saturated zone (Rempe & Dietrich, 2018; Salve et al., 2012).

How should rock moisture dynamics be incorporated into simple catchment modelling frameworks? Does the incorporation of this deep vadose zone storage dynamic as a modelling target reduce uncertainty and improve accuracy in other model parameters and outputs? Do tradeoffs in model performance emerge when both storage dynamics and discharge are constrained? Here, we approach these questions by confronting a lumped catchment model with observations of both discharge and rock moisture dynamics. Rock moisture dynamics are obtained from a unique vadose zone monitoring system consisting of an inclined borehole with nine moisture sensors distributed throughout a 14 m thick weathered bedrock vadose zone. The intensively monitored hillslope drains to the Elder Creek catchment in the seasonally dry Northern California Coast Ranges. We formulate a hydrological model based on existing conceptual understanding of runoff generation and storage at the site. We find that when the model is fit to discharge alone, storage parameters are poorly constrained and in general inaccurately describe storage dynamics. In contrast, when the model is fit to both storage and discharge, model error with respect to discharge remains low, yet model results (in terms of accurately simulating rock moisture and constraining the post calibration parameter space) are significantly improved.

2 SITE DESCRIPTION

Elder Creek, the study catchment where streamflow is measured and modelled, is part of the Eel River Critical Zone Observatory (ERCZO) and lies approximately 160 km north of San Francisco in the Northern California Coast Ranges. Elder Creek (catchment area 16.9 km^2) is a tributary to the South Fork Eel River. Discharge has been monitored approximately 0.6 km upstream from the catchment mouth continuously since October, 1967, by the USGS ($https://waterdata.usgs.gov/nwis/uv?site_no = 11475560$). Elder Creek is part of the USGS Hydrological Benchmark Network, a collection of long-term monitoring watersheds with minimal anthropogenic impact (Mast & Clow, 2000).

Rivendell, the study hillslope where rock moisture is measured, is 0.3 km upstream from the mouth of Elder Creek, and has been the site of intensive hydrological (Oshun et al., 2016; Rempe & Dietrich, 2018; Salve et al., 2012; Schmidt & Rempe, 2020; Vrettas & Fung, 2015), ecological (Hahm, Rempe, et al., 2019; Link et al., 2014; Simonin et al., 2014; Tune et al., 2020), and geochemical (Bilir et al., 2021; Druhan et al., 2017; Kim et al., 2017) monitoring. Rivendell and the lower portion of Elder Creek are situated on what is currently the University of California's Angelo Coast Range Reserve; the headwaters of Elder Creek are Bureau of Land Management Wilderness area.

2.1 Climate

The study area experiences a Mediterranean climate (Peel et al., 2007) with hot, dry summers and cool, wet winters with intense precipitation. On average about 2000 mm of precipitation falls between the months of November and May, with little to no precipitation during the dry season (Daly et al., 2015). Snowfall is rare and melts a few days after falling. The mean annual temperature is 12°C (Daly et al., 2008).

2.2 Vegetation

The plant community within the Elder Creek catchment is primarily mixed broadleaf-needleleaf evergreen forest, with some chaparral on south-facing, higher elevation slopes and some deciduous hardwoods along the riparian zone. The watershed contains part of the largest remaining old-growth Douglas fir (*Pseudotsuga menziesii*) forest in California. Large, canopy-dominant Douglas fir tend to be concentrated on north facing slopes and topographically convergent areas. Other abundant trees in the watershed include madrone (*Arbutus menziesii*), bay laurel (*Umbellularia californica*), live oaks (*Quercus chrysolepis* and *agrifolia*), and tan oak (*Notholithocarpus densiflorus*) (Hahm, Rempe, et al., 2019). Trees in the riparian zone include big leaf maple (*A. macrophyllum*) and alder (*Alnus*).

2.3 Physiography

The catchment is steep, with a mean hillslope gradient of about 50% (Lovill et al., 2018), and ranges from 385 to 1285 m above sea level. Rivendell is a steep (approximately 62%), unchanneled, north-facing hillslope (Lidar-derived hillshade is shown in Figure 1) (Oshun et al., 2016).

2.4 Bedrock, soils

The watershed is underlain by steeply dipping deformed turbidite sequences of the Coastal Belt of the Franciscan Complex (Ernst & McLaughlin, 2012; Jayko et al., 1989), consisting of interbedded argillite (mudstone) and greywacke (sandstone), with minor conglomerate (Hahm, Rempe, et al., 2019; Lovill et al., 2018). Hillslopes tend to be soil-mantled, with some exposed rock outcrops on ridges, especially those underlain by sandstone (Lovill et al., 2018; Oshun et al., 2016; Rempe & Dietrich, 2018). Soil thickness is locally variable but typically thin (0.5 m). Beneath the mobile soil lies an oxidized, highly fractured saprolite (up to 4 m thick) that retains relict bedrock structure (e.g., bedding planes), that transitions with depth to a progressively less oxidized, fractured bedrock (Salve et al., 2012). Deep drilling has revealed that the thickness of the weathered bedrock zone increases from channel to the hillslope divide (Oshun et al., 2016; Rempe & Dietrich, 2014; Salve et al., 2012). In the channel, fresh bedrock is exposed in the bed, whereas at the ridgetop it can be 25–30 m deep (see Figure 2).

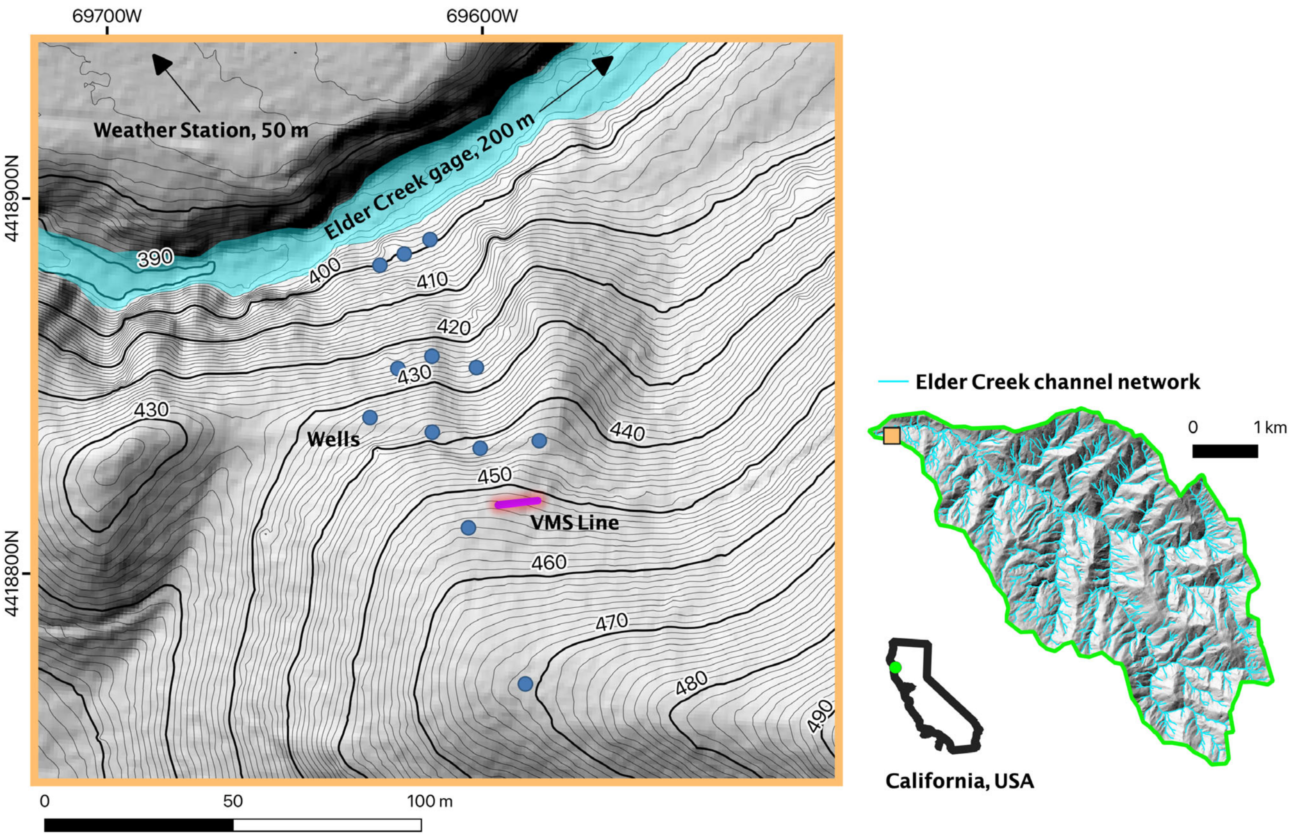

FIGURE 1 The Rivendell experimental hillslope (left) with coordinates in NAD83/UTM zone 10 N (EPSG:26910). Located at approximately 455 m elevation, the vadose zone monitoring system (VMS) is illustrated with a neon purple line. Blue dots are well locations where rock moisture is surveyed via neutron probe. The Elder Creek catchment boundary (right) outlined in green with channel network in blue, based on Lovill et al. (2018). The Rivendell hillslope location is denoted with an orange point

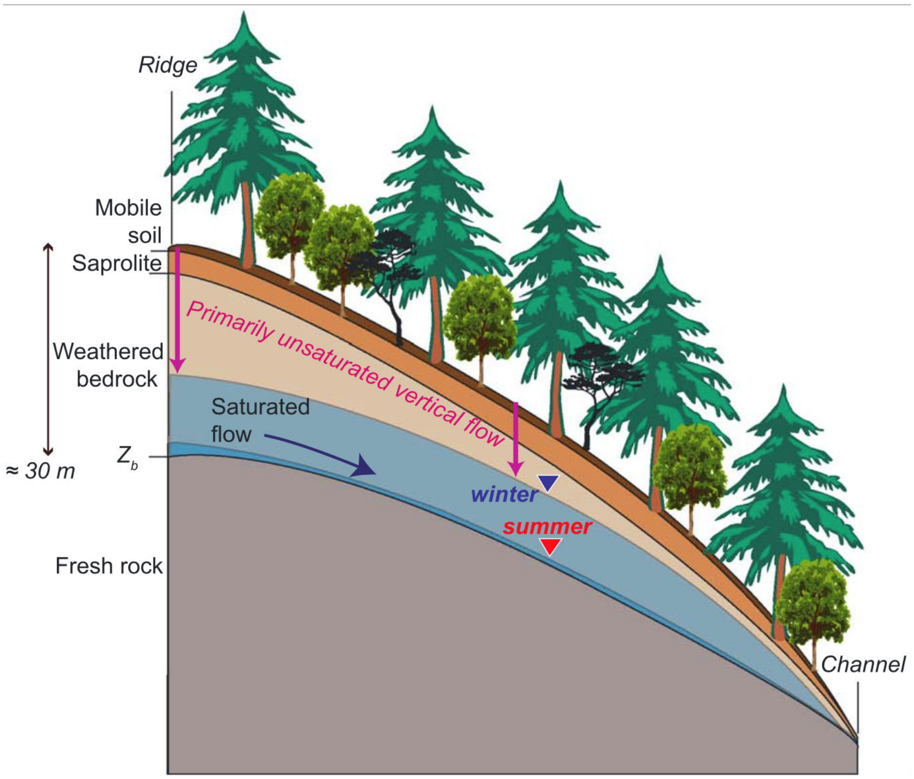

FIGURE 2 Conceptual cross-section of the Rivendell hillslope, illustrating subsurface weathering profile and hydrologic zones and fluxes. Soil moisture occurs within the thin mobile soil; rock moisture occurs within the saprolite and weathered bedrock. Seasonal groundwater (blue shades) occurs within the weathered bedrock above Z_b, which marks the boundary between weathered and fresh bedrock. Inverted triangles denote water table positions. Figure adapted from Hahm, Rempe, et al. (2019)

2.5 Conventional hydrologic modelling data

The United States Geological Survey (USGS) Elder Creek gaging station (gauge ID: 11475560) provides estimates of stream discharge and is located about 200 m upstream from the instrumented hillslope. Hydroclimatic forcing data for the model is derived from a weather station approximately 0.5 km from the USGS gauge; this weather station also provides air temperature. Precipitation is derived from a Campbell Scientific Model TB4 tipping bucket rain gague. A correction factor is applied to the rainfall time series to account for wind-induced undercatch (Yang et al., 1998).

Potential evapotranspiration values are obtained with the Hargreaves equation (Hargreaves & Samani, 1985):

$$\mathrm{PET} = 0.0023 \cdot (T_{\mathrm{mean}} + 17.8) \cdot (T_{\mathrm{max}} - T_{\mathrm{min}})^{0.5} \cdot 0.408 \cdot R_{\mathrm{ext}} \quad (1)$$

Here, PET is potential evapotranspiration (mm/d), T_{mean} is the mean daily temperature (°C), T_{max} and T_{min} are respectively the maximum and minimum daily temperatures (°C), and R_{ext} is the extraterrestrial solar radiation (W/m^2), calculated using latitude and the day of year via the procedure used by Allen et al. (1998). After the daily PET is calculated, it is resampled to the hourly time step, such that the PET is the same for each hour of a given day. Finally, all data are available at the 15 min or smaller temporal resolution; after PET is calculated, all forcing or calibration data are resampled to the hourly time step.

The Hargreaves formulation is capable of giving PET values without considering interception. However, because the Elder Creek catchment contains significant vegetation, a procedure for estimating interception should be implemented. To achieve this, the procedure employed by Dralle et al. (2018) is used, where a value of 4 mm/d is subtracted from all precipitation intensities (unless doing so would reduce the precipitation intensity to a value less than 0 mm/d, in which case the precipitation intensity is set to 0 mm/d). Because we employ the hourly temporal resolution for forcing data, we subtract up to 0.1667 mm/h from each step of the hourly undercatch corrected precipitation data. All forcing or calibration time series are shown for the entire modelling period in Figure 3.

2.6 Rock moisture data

Two unique observational capabilities at the Elder Creek watershed exist to capture the dynamics of rock moisture storage. First, boreholes that span the unsaturated zone from channel to ridgetop enable direct quantification of gains and losses in unsaturated weathered bedrock (rock moisture) storage via monthly neutron probe surveys (Rempe & Dietrich, 2018; Salve et al., 2012) and nuclear magnetic resonance campaigns (Schmidt & Rempe, 2020). Second, an inclined borehole that spans the weathered bedrock unsaturated zone is outfitted with distributed time domain transmission (TDT) sensors for

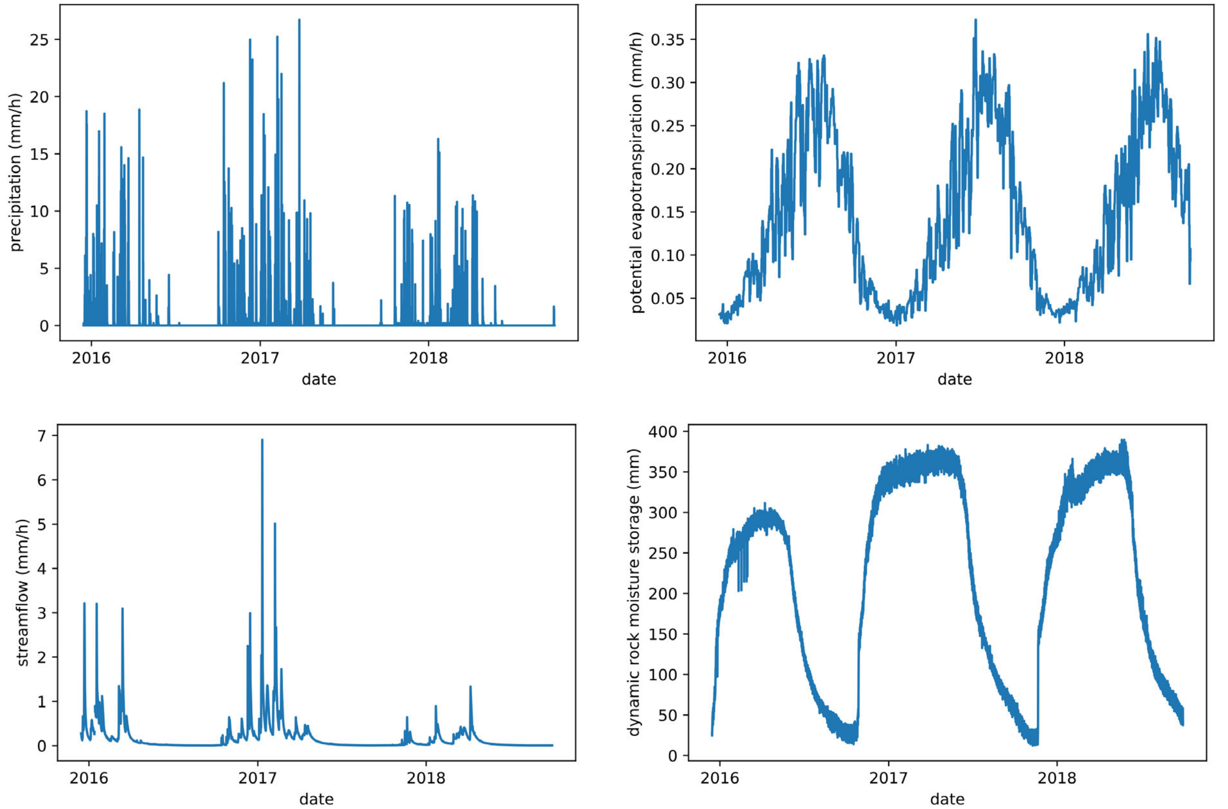

FIGURE 3 Hourly forcing and calibration data from the Elder Creek catchment used by the model. The entire modelling period is shown, including the 0.8 year spin up period (beginning on 15 December 2015), the calibration period (water year 2017), and the validation period (water year 2018). Upper left: Precipitation data, corrected for interception and undercatch. Upper right: Potential evapotranspiration data, calculated with the Hargreaves equation. Lower left: Streamflow data. Lower right: Catchment average dynamic rock moisture data

continuous monitoring of rock moisture dynamics as part of a Vadose-zone Monitoring System (VMS) (Tune et al., 2020). The location of the VMS is shown in Figure 1. The hillslope where the monitoring boreholes and VMS are located (named Rivendell) is broadly representative of the Elder Creek catchment as a whole, which is similarly composed of steep hillslopes bounded by debris-flow incised channels, overlain by a dense old-growth plant community cover and underlain by steely dipping turbidites (Hahm, Rempe, et al., 2019; Lovill et al., 2018; Seidl et al., 1992). Here, we combine the quantitative storage estimates from neutron probe surveys with the continuous time series from the TDT sensors to arrive at a representative rock moisture storage time series for the Elder Creek catchment. To accomplish this, we first averaged the TDT time series for sensors located in the unsaturated zone year-round (1.95, 3.64 and 5.22 m depth) to capture the temporal pattern of storage accumulation and depletion. We then normalized this relative time series using the range of storage representative of the Rivendell hillslope (0–400 mm annual storage) (Schmidt & Rempe, 2020), which is measured via neutron probe in boreholes.

3 METHODS

For some catchments, the apparent importance of rock moisture's role in providing water for plants and mediating recharge to groundwater (and therefore generating streamflow) motivates inclusion of rock moisture in hydrologic models. We develop a model that adheres to observed dynamics within Elder Creek, and enables multi-criteria comparison of model outputs to unique datasets available within the Elder Creek watershed, including rock moisture and streamflow observations. In light of the hydrologic relevance of rock moisture, MCA on both streamflow and rock moisture is a promising technique which might reduce parameter uncertainty and improve accuracy beyond streamflow in modelling efforts.

3.1 Pertinent hydrologic observations from previous studies that guide model development

Previous studies have documented subsurface weathering profile structure and associated water storage and runoff dynamics at the Elder Creek watershed (Dralle et al., 2018; Hahm, Rempe, et al., 2019; Lovill et al., 2018; Oshun et al., 2016; Rempe & Dietrich, 2014, 2018; Salve et al., 2012). Based on detailed measurements (see Figure 1, Rempe & Dietrich, 2018), Figure 2 depicts what we propose to be a representative hillslope profile, with a relatively thin soil (typically < 1 m), underlain by a weathered bedrock layer that increases in thickness from channel (approximately 0–2 m thick) to ridge (up to 30 m thick). Plant water use throughout the long, dry growing season (May–October) generates large storage deficits in the (unsaturated) root zone, at which point early wet season rainfall (October–December) wets thin soils but also penetrates into the underlying saprolite (Salve et al., 2012). Eventually in the wet season the soil moisture climbs to an annual consistent maximum value of about 100 mm of unsaturated storage (Rempe & Dietrich, 2018). Rainfall that infiltrates beyond the soil wets up into unsaturated weathered bedrock (which includes the upper most saprolite), which (much like the soils) consistently wets to a seasonal maximum. However, maximum observed dynamic water storage in the bedrock vadose zone ranges from 100 mm near the channel to upwards of 600 mm near the ridge, due to the great depth of the weathered bedrock. Following the seasonal wetting of the unsaturated zone, primarily vertical flow recharges the end of dry season groundwater developed in the weathered bedrock above the fresh bedrock. Successive rainfall events then leads to rapid rise of the groundwater by many meters. This drives lateral transport through a network of fractures that release water to channels primarily through discrete seeps and springs. Although the weathered bedrock vadose zone is thick, groundwater—and thus streamflow—response to precipitation is rapid (on the order of hours) (Salve et al., 2012). Notably, changes in saturated storage in the catchment directly impact streamflow (Dralle et al., 2018), and overland flow is rare on hillslopes and has only been observed as local saturation overland flow in some unchanneled valleys during exceptional rainfall events. at Elder Creek.

3.2 Model description

We constructed a lumped rainfall-runoff model which captures many of the relevant hydrological processes described in Sections 2 and 3.1. Given forcing data and state variables for a time step, the model

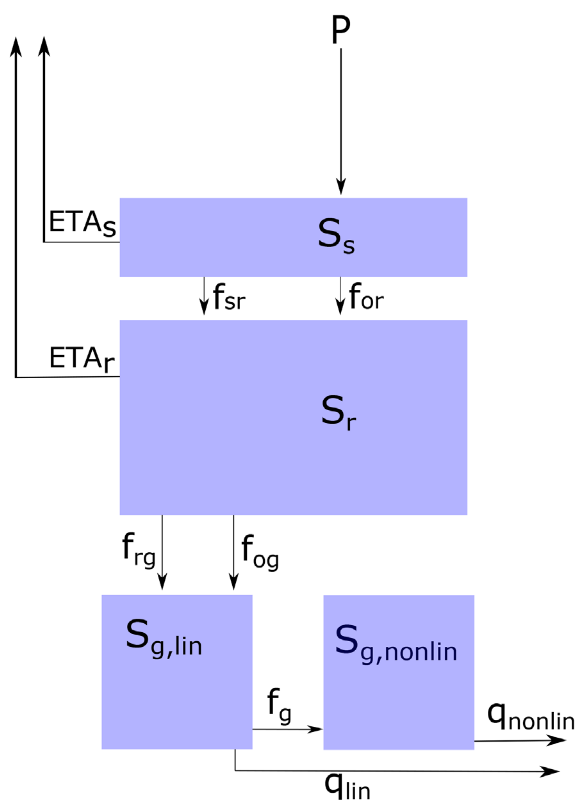

FIGURE 4 Elder Creek model wiring diagram

simulates its state variables for the next time step via calculating the flux in each state variable for the given time step. All state variables, fluxes, and individual components of fluxes are recorded. A detailed description of the model can be found in Figure 4, Tables 1–3, and Equations (2)–(14).

The model is forced by precipitation and potential evapotranspiration. It has four state variables, all with units of mm: S_s, the soil moisture storage; S_r, the rock moisture storage (which is located in the bedrock vadose zone); $S_{g,lin}$, the linear groundwater reservoir storage, and $S_{g,nonlin}$, the nonlinear groundwater reservoir storage. Precipitation first enters the soil moisture storage, which then leaks water to the bedrock vadose zone as a power law, conceptually similar to gravity-driven leakage controlled by hydraulic conductivity as a function of relative soil moisture (see Equation (4)). If soil moisture would exceed its maximum value, all excess (overflow) water is directly sent to the bedrock vadose zone (see Equation (5)). Note that this formulation allows for a seasonal maximum in soil moisture, as well as for small amounts of leakage to the bedrock vadose zone before the soil moisture is at its seasonal maximum. Similarly, drainage from rock moisture to the groundwater occurs in two ways: gravity drainage and overflow drainage. Gravity drainage is simulated as a power law, using the parameters k_{sat}, as the saturated hydraulic conductivity of the fractured rock, and b_{fc}, the power to which the relative rock moisture $S_r/S_{r,max}$ is raised in order to determine unsaturated hydraulic conductivity as a fraction of saturated hydraulic conductivity. Overflow drainage from the rock moisture drains all water that would cause the rock moisture to exceed a given threshold $S_{r,max}$ (see Equation (9)). Because the vadose zone is thick year-round (and increases in thickness during summer groundwater drainage to streams), we assume that roots do not reach the groundwater and so evapotranspiration only comes from the two vadose state variables over the majority of the catchment. The potential evapotranspiration that the soil and bedrock vadoze zone experience is controlled by the root weighting parameter r, where a value of 1 indicates that all roots (and therefore all evapotranspirative demand) are in the soil, a value of 0.5 indicates equal root distribution between the soil and bedrock vadose zone, and a value of 0 indicates that all roots are in the bedrock vadose zone. Potential evapotranspiration is further corrected to actual evapotranspiration for each vadose state variable, as a function of how full each vadose state variable is above the wilting point parameter (see Equations (3) and (7)).

TABLE 1 Elder Creek model state variables. All state variables have units of mm

State variable	Description
S_s	Soil moisture storage
S_r	Rock moisture storage
$S_{g,lin}$	Linear groundwater reservoir storage
$S_{g,nonlin}$	Nonlinear groundwater reservoir storage

TABLE 2 Elder Creek model fluxes. All fluxes have units of mm/h

Flux	Description	Type
P	Precipitation	External
ETP	Total potential evapotranspiration	External
ETA_s	Actual evapotranspiration from the soil	External
ETA_r	Actual evapotranspiration from the bedrock vadose zone	External
q_{lin}	Discharge from the linear groundwater reservoir	External
q_{nonlin}	Discharge from the nonlinear groundwater reservoir	External
f_{sr}	Gravity drainage from soil to bedrock vadose zone	Internal
f_{or}	Overflow leakage from soil to bedrock vadose zone	Internal
f_{rg}	Gravity drainage from rock moisture to groundwater	Internal
f_{og}	Overflow drainage from rock moisture to groundwater	Internal
f_g	Flow from linear to nonlinear groundwater reservoir	Internal

Two groundwater reservoirs are used to simulate saturated storage; discharge to the stream is only possible from these reservoirs, consistent with the observation that streamflow only comes from saturated storage in Elder Creek (Dralle et al., 2018; Lovill et al., 2018; Rempe & Dietrich, 2018; Salve et al., 2012). We found the inclusion of two reservoirs to be necessary to adequately represent recession limbs; when a single groundwater reservoir was used, KGEs in streamflow rarely exceeded 0.7. Although it was not explicitly our intention, the inclusion of two groundwater reservoirs can be conceptualized as a form of semi distribution (Harman et al., 2009). Drainage from the vadose zone first enters the linear groundwater reservoir, which can directly produce discharge or send water to the nonlinear reservoir. The nonlinear reservoir is also capable of contributing to discharge.

$$\frac{\Delta S_s}{\Delta t} = P - ETA_s - f_{sr} - f_{or} \tag{2}$$

$$ETA_s = \max\left(0, \frac{S_s - S_{s,max} \cdot s_{wilt}}{S_{s,max}} \cdot \frac{1}{1 - s_{wilt}} \cdot r \cdot ETP\right) \tag{3}$$

$$f_{sr} = k_{s,sat} \cdot (S_s/S_{s,max})^{b_s} \tag{4}$$

$$f_{or} = \begin{cases} (S_s + (P - ETA_s - f_{sr}) \cdot \Delta t - S_{s,max})/\Delta t \\ \quad \text{if } S_s + (P - ETA_s - f_{sr}) \cdot \Delta t > S_{s,max} \\ 0 \\ \quad \text{if } S_s + (P - ETA_s - f_{sr}) \cdot \Delta t \leq S_{s,max} \end{cases} \tag{5}$$

$$\frac{\Delta S_r}{\Delta t} = f_{sr} + f_{or} - f_{rg} - f_{og} - ETA_r \tag{6}$$

$$ETA_r = \max\left(0, \frac{S_r - S_{r,max} \cdot s_{wilt}}{S_{r,max}} \cdot \frac{1}{1 - s_{wilt}} \cdot (1 - r) \cdot ETP\right) \tag{7}$$

TABLE 3 Elder creek model parameters. Parameters ranges were chosen to exceed physically realistic values, such that a physically realistic parameter space can be covered in calibration. The exception is $S_{s,max}$, which is constrained to a value of 100 mm and is therefore not included in calibration

Parameter	Units	Description	Calibration range
r	--	Relative fraction of roots in soil layer; controls distribution of ET between soil and bedrock vadose zone	0.001–1
$S_{s,max}$	mm	Maximum soil water storage; controls threshold at which soil leakage begins, as well as correction of ET	100
$S_{r,max}$	mm	Maximum rock moisture storage; controls threshold at which overflow drainage to groundwater begins, rate of gravity drainage from rock moisture to groundwater, and correction of ET	500–20 000
s_{wilt}	–	Relative moisture of $S_{r,max}$ and $S_{s,max}$ below which no actual evapotranspiration occurs	0–0.5
b_s	–	Power controlling shape of gravity drainage from soil moisture to rock moisture	1–30
$k_{s,sat}$	mm h^{-1}	Hydraulic conductivity of soil when soil moisture is at its seasonal maximum	0.01–1000
b_r	–	Power controlling shape of gravity drainage from rock moisture to groundwater	1–40
$k_{r,sat}$	mm h^{-1}	Saturated hydraulic conductivity of fractured rock, which controls rate of gravity drainage to groundwater	4–1000
a	(mm)$^{1-b}$ h^{-1}	Coefficient for nonlinear groundwater reservoir	$5 \cdot 10^{-5}$–0.125
b	–	Exponent for nonlinear groundwater reservoir	0.5–3
k_1	h^{-1}	Coefficient which controls the discharge of the linear groundwater reservoir	$5 \cdot 10^{-5}$–0.125
k_{12}	h^{-1}	Linear groundwater reservoir coefficient which controls the rate at which water flows from the linear saturated store to the nonlinear saturated store	$5 \cdot 10^{-5}$–0.125

$$f_{rg} = k_{r,sat} \cdot (S_r / S_{r,max})^{b_r} \tag{8}$$

$$f_{og} = \begin{cases} (S_r + (f_{sr} + f_{or} - ETA_r - f_{rg}) \cdot \Delta t - S_{r,max}) / \Delta t \\ \quad \text{if } S_r + (f_{sr} + f_{or} - ETA_r - f_{rg}) \cdot \Delta t > S_{r,max} \\ 0 \\ \quad \text{if } S_r + (f_{sr} + f_{or} - ETA_r - f_{rg}) \cdot \Delta t \le S_{r,max} \end{cases} \tag{9}$$

$$\frac{\Delta S_{g,lin}}{\Delta t} = f_{og} + f_{rg} - q_{lin} - f_g \tag{10}$$

$$q_{lin} = k_1 \cdot S_{g,lin} \tag{11}$$

$$f_g = k_{12} \cdot S_{g,lin} \tag{12}$$

$$\frac{\Delta S_{g,nonlin}}{\Delta t} = f_g - q_{nonlin} \tag{13}$$

$$q_{nonlin} = a \cdot (S_{g,nonlin})^b \tag{14}$$

3.3 Numerical method

The model functions via calculating the flux for each state variable for a given time step. Therefore, this model is composed of a system of four ordinary nonlinear differential equations (one equation describing the net flux in each state variable; see Equations (2), (6), (10), and (13)). This coupled system of nonlinear equations cannot be solved analytically, so we must rely on a numerical method which approximates the exact solutions of these equations. We employ the explicit adaptive midpoint (second-order) method to do this. Explicit adaptive second-order methods have been shown to be robust choices in the context of lumped rainfall-runoff models (Clark & Kavetski, 2010; Kavetski & Clark, 2010; La Follette et al., 2021; Schoups et al., 2010). This method has an adaptive time step: numerical error is automatically controlled by shortening the time step when numerical error is too large, and the time step is relaxed when numerical error is small for faster computation. Details on this algorithm can be found in Appendix A.

3.4 Description of analysis

Nearly 3 years of data are used (from 15 December 2015 to 1 October 2018). The first 0.8 years (from 15 December 2015 to 30 September 2016) are used as a spin up period. Water year 2017 is used as a calibration period, and then water year 2018 is used for validation (to provide further confidence that the optimal parameter sets generated for water year 2017 are appropriate for the catchment rather than fit to the calibration period in particular). At the start of the spin up period, initial conditions in soil moisture (Oshun et al., 2016) and rock moisture are set using observed values at the beginning of the spin up period. Note that the soil moisture datum is the average normalized value given by two TDR moisture probes located at 15 and 35 cm depth in the soil at the ridgetop near well 15, at the start of 15 December 2015, and soil moisture data are not used aside from setting initial conditions here. Specifically, soil moisture was observed to be at 51% of its maximum value at the first time step, and rock moisture to be 372.7 mm below its maximum value; therefore, the initial condition in soil moisture is set to 51% of $S_{s,max}$ and the initial rock moisture is set to $S_{r,max}$ – 372.7 mm. Initial conditions in the saturated state variables are arbitrary and work their way out of the system quickly because the catchment is flashy and discharge comes solely from saturated storage.

Parameters sets are generated via a Latin hypercube approach. In total, 10 000 parameter sets are generated. We determine that this is a sufficient number of parameter sets due to the fact that the results of our study do not change when repeated with a different selection of 10 000 parameter sets. Note that the model has, in total, 12 parameters, and of these, 11 are included in calibration. The parameter $S_{s,max}$ is fixed to a value of 100 mm and is not included in calibration. This value is selected to reflect observations which show that the maximum soil moisture of the catchment tends to be 100 mm (Rempe & Dietrich, 2018). Each parameter set is used to make a prediction of streamflow and rock moisture over the entire time period. While the model calculates total rock moisture, the observations of rock moisture are dynamic rather than total—that is, observations in rock moisture are equal to the amount of moisture above some minimum value during the total modelling period. Therefore, the model calculates dynamic rock moisture by subtracting the minimum simulated rock moisture during the total simulation period from the simulated rock moisture time series.

Next, the Kling-Gupta efficiency, or KGE (Gupta et al., 2009) is calculated for both the streamflow and the dynamic rock moisture resulting from each parameter set during the calibration period. The KGE is a way to assess how closely one time series matches another, equally weighting components of bias, variance, and correlation. Note that KGE values above -0.41 represent simulations that perform better than the mean value in observed data (Knoben et al., 2019) and a KGE value of 1 indicates a perfect simulation. For each parameter set, the KGE in simulated dynamic rock moisture is plotted against the KGE in simulated streamflow. This forms a 2D KGE scatterplot, which allows us to examine the trade-off in model performance in rock moisture and streamflow. Then, to determine the sensitivity of optimal parameters to calibration strategy, three calibration scenarios are considered: optimized results in streamflow only, optimized results in rock moisture only, and Pareto optimized results in both streamflow and rock moisture. Optimal parameter sets for each calibration scenario are considered to be the sets which yield performance within the top 1% of all parameter sets for the given optimization metric. Performance in streamflow or dynamic rock moisture alone is simply calculated via KGE in either variable. The Pareto score is defined as $\sqrt{(\mathrm{KGE_{RM}}-1)^2+(\mathrm{KGE_q}-1)^2}$; therefore Pareto optimality is achieved as this score is minimized, or in other words when the KGEs in dynamic rock moisture and streamflow both approach 1. The Pareto score in this manuscript is equivalent to the distance from a perfect simulation in Pareto space; achieving Pareto optimality via minimizing Euclidean distance from an optimal point has seen previous use in hydrologic modelling (Gupta et al., 2009; Koppa et al., 2019). Distributions in parameters within optimal parameter sets were compared between the three calibration scenarios, in order to determine to what extent calibration scenario affects optimal parameter values. Average performance metrics for the three different calibration scenarios are reported for the calibration and validation periods, in order to assess the extent to which optimal parameters represent the catchment itself (as opposed to overfitting to a certain time period). Splitting data into calibration and validation periods has been shown to be beneficial in other modelling efforts which simulate more hydrologically relevant variables than just streamflow (Feyen et al., 2000; Refsgaard, 1997).

4 RESULTS

Figure 5 shows, in the left-hand column, storage states for soil, rock moisture, and groundwater reservoirs, and in the right-hand column, fluxes out of these reservoirs for the top 1% performing parameter sets calibrated on both dynamic rock moisture and streamflow. Dynamic rock moisture has the highest seasonal dynamic range (> 500 mm), followed by soil moisture (< 100 mm), with the combined groundwater reservoirs exhibiting a relatively small (< 150 mm) range in the 2017 water year. Modelled soil moisture rises quickly, then plateaus for most of the wet season, with high magnitude drainage pulses that correspond to incoming precipitation events (although some drainage occurs when soil moisture is below its maximum value). Modelled rock moisture rises more slowly, and exhibits larger recessions and rises within the wet season. Increases in modelled groundwater lag behind soil and rock moisture, and are relatively muted. All modelled drainage fluxes from these reservoirs, when they occur, are much larger in magnitude (and more comparable to precipitation intensities) than modelled ET, which is constrained by PET, which is shown in Figure 6. Figure 6 also shows how modelled ET can be preferentially drawn from soil moisture in the wet season, but has an increasing (and eventually larger) rock moisture component over the course of the dry season.

From Figure 7, it is apparent that some parameter sets used by the model are capable of providing relatively high accuracy results in streamflow or dynamic rock moisture, where the maximum KGE obtained in streamflow simulation is equal to 0.91 and the maximum KGE obtained in dynamic rock moisture simulation is equal to 0.92. It

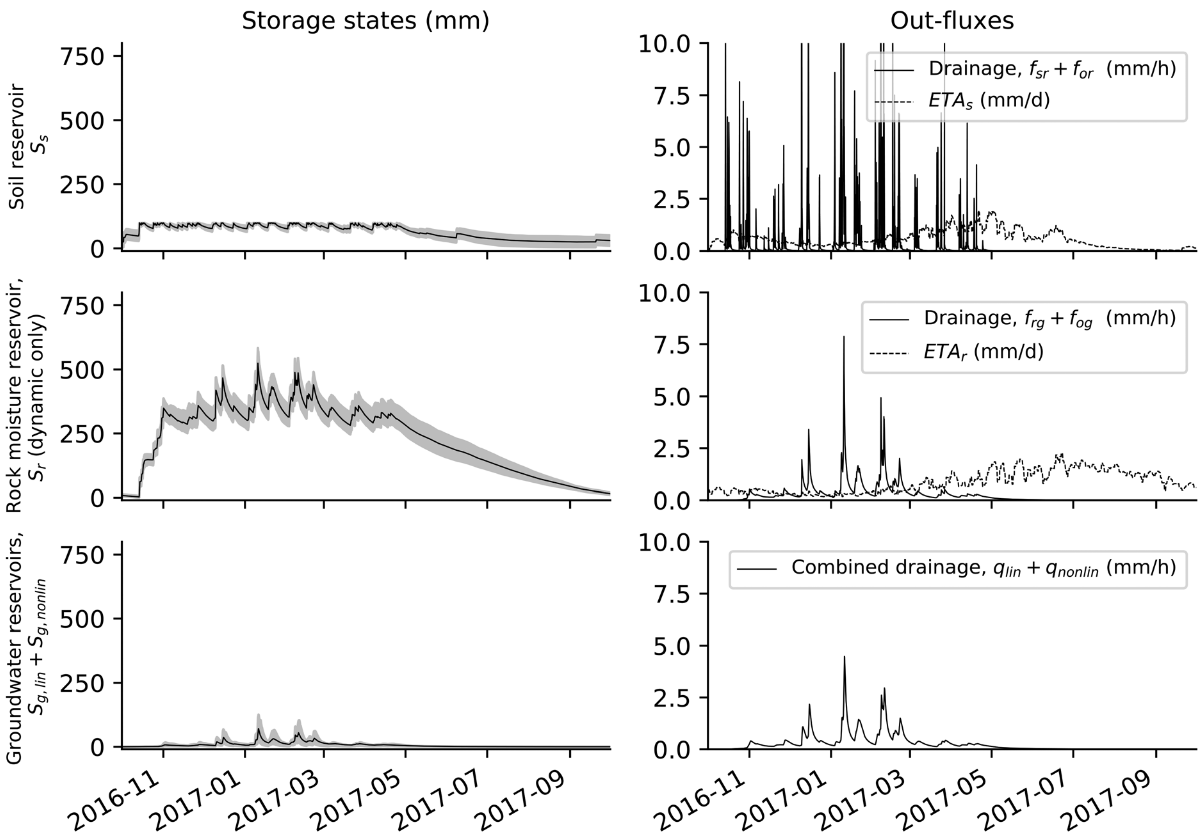

FIGURE 5 Modelled storage states (left-hand column) and out-fluxes (right-hand column) of the soil, rock moisture, and groundwater reservoirs (first, second, and third rows, respectively), for the 2017 water year. The soil and groundwater reservoir storages are expressed in terms of their actual magnitude, whereas the rock moisture reservoir storage is expressed in terms of its dynamic range (i.e., relative to its lowest achieved value). All lines represent the mean of the top 1% performing models simultaneously calibrated on both rock moisture and streamflow. Shaded intervals for storage states cover the 10th to 90th percentiles of the top 1% performing parameter sets. Solid out-flux lines denote drainage fluxes (in mm/h, from the soil reservoir to rock moisture, from rock moisture to groundwater, and from both groundwater reservoirs to streamflow across rows, respectively), and dotted lines denote ET fluxes (in mm/d)

is also apparent that some parameter sets are capable of simultaneously simulating both streamflow and dynamic rock moisture accurately, where the parameter set with the highest Pareto optimality yields KGEs in streamflow and dynamic rock moisture simulations of 0.87 and 0.81, respectively. Meanwhile, calibrating on streamflow only can yield good results in simulated streamflow but poor results in simulated rock moisture, where the parameter set that yielded optimal performance in streamflow simulation yielded a KGE of 0.53 in dynamic rock moisture simulation. More generally, it is clear that the top 1% optimal parameter sets in streamflow simulation (blue points in Figure 7) do not necessarily yield high KGEs in simulated dynamic rock moisture. Similarly, the top 1% optimal parameter sets in dynamic rock moisture simulation (green points in Figure 7) do not necessarily yield high KGEs in simulated streamflow. Therefore, Pareto optimization on both dynamic rock moisture and streamflow offers improved model results in terms of simultaneously simulating rock moisture and streamflow when compared to calibration scenarios which optimize on a single variable, with minimal sacrifice in simulation quality for streamflow or dynamic rock moisture individually. Figure 8 provides visual perspective on how simulations compare to observations for the 2017 water year.

Next, we examined the distributions of parameters resulting from the top 1% of each calibration scenario to determine to what extent optimal parameter values are impacted by choice of calibration scenario. We found that optimal parameter distributions were usually insensitive to calibration scenario, indicated by the fact that parameter distributions after calibration either spanned the entire calibration range for all calibration scenarios, the average calibrated parameter values changed very little between different calibration scenarios, or calibrated parameter ranges were very similar between different calibration scenarios. These observations suggest heavy equifinality within the model. Notably, however, two parameters did not exhibit this behaviour: r and $k_{s,sat}$. Distributions of these parameters are shown in Figure 9, demonstrating potential effects of calibration scenario on optimal parameter values.

From Figure 9, left, it is apparent that the parameter r, which controls the relative distribution of roots (and therefore evapotranspirative demand) between the soil and bedrock vadose zone, can nearly span its entire possible range and still yield a high quality simulation in streamflow. In contrast, optimal r values in the dynamic rock moisture-only calibration scenario are restricted to a smaller subset of the parameter space or r, where this restriction

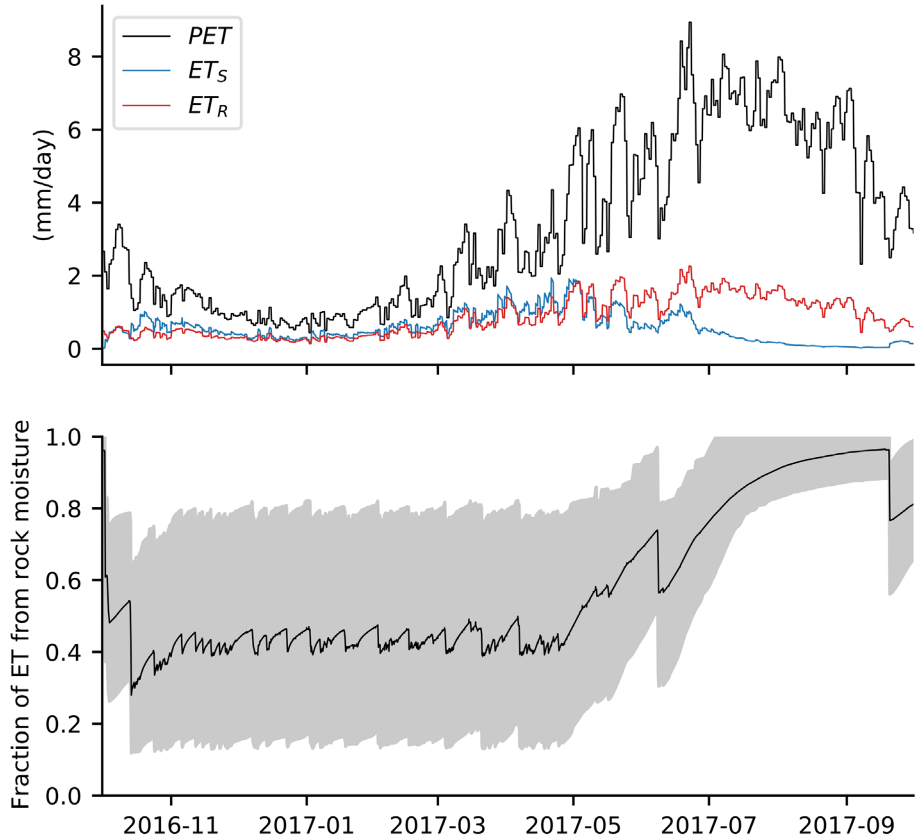

FIGURE 6 (top) PET model forcing, and mean modelled ET from soil (ET_S) and rock moisture (ET_R) from the top 1% performing models calibrated simultaneously on both rock moisture and streamflow. (bottom) mean (solid line) and 10th–90th percentile (shaded region) fraction of ET sourced from rock moisture from the top 1% performing models calibrated simultaneously on both rock moisture and streamflow

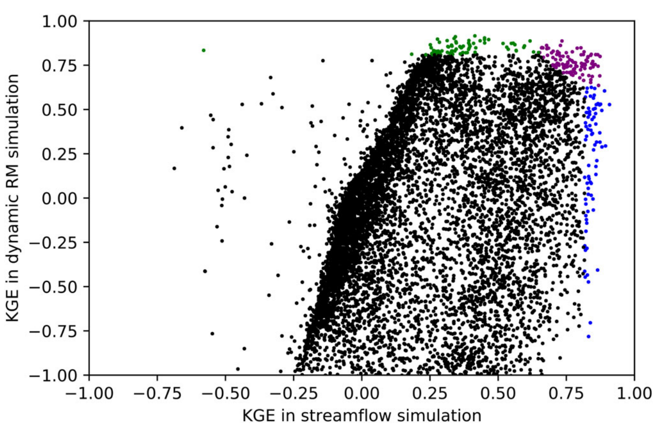

FIGURE 7 2D KGE scatterplot, showing the KGEs in streamflow and dynamic rock moisture simulations resulting from 10 000 parameter sets. Blue points indicate parameter sets which performed within the top 1% in terms of KGE in streamflow simulation, green points indicate parameter sets which performed within the top 1% in terms of KGE in dynamic rock moisture simulation, and purple points indicate parameter sets which performed within the top 1% in terms of Pareto optimality. Note that if a parameter set achieves performance within the top 1% of dynamic rock moisture or streamflow simulations individually but also performed within the top 1% in terms of Pareto optimality, it will appear as purple

relaxes somewhat in case of Pareto optimization. Because previous work in Elder Creek has shown that roots are distributed between both the soil and bedrock vadose zone (Hahm, Rempe, et al., 2019; Rempe & Dietrich, 2018; Schmidt & Rempe, 2020; Tune et al., 2020), it is clear that calibration on streamflow alone can yield values in r that are very close to 1, which we believe to be physically unrealistic.

In the right plot of Figure 9, distributions in the parameter $k_{s,sat}$ can be seen. As a broad variety of saturated hydraulic conductivites can be found throughout the soils of the catchment, we do not comment on the physical realism of these optima; rather, we merely report that $k_{s,sat}$ is clearly sensitive to choice of calibration target.

For the top 1% of all parameter sets resulting from each calibration scenario, KGEs were evaluated for both the calibration and validation periods. Specifically, we calculated KGEs in streamflow during calibration and validation of parameter sets that achieve within the top 1% performance of KGE in streamflow in the calibration period, KGEs in dynamic rock moisture during calibration and validation of parameter sets that achieve within the top 1% performance of KGE in dynamic rock moisture in the calibration period, and both KGEs in streamflow and dynamic rock moisture during calibration and validation of parameter sets that achieve within the top 1% performance of Pareto optimality during the calibration period. Mean values of these KGEs, and the difference in means between calibration and validation periods, are reported here. The decrease in KGE mean values for streamflow in the streamflow calibration scenario, between calibration and validation periods, is 0.10 (from 0.843 to 0.739); the decrease in KGE mean values for rock moisture in the rock moisture calibration scenario, between calibration and validation periods, is 0.03 (from 0.844 to 0.813); the decrease in KGE mean values for streamflow in the Pareto calibration scenario, between calibration and validation

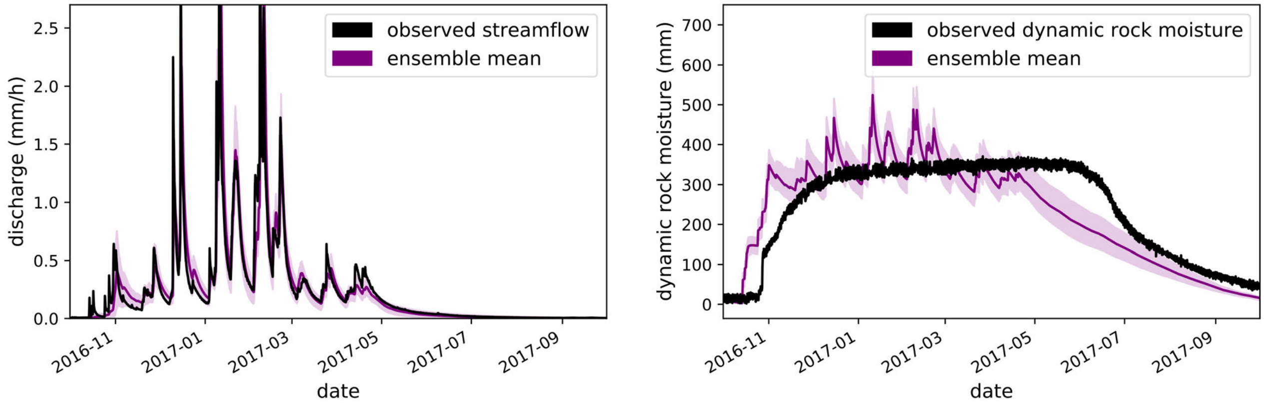

FIGURE 8 Ensemble mean simulation of streamflow (left) and dynamic rock moisture (right) resulting from the top 1% of all parameter sets in terms of Pareto optimality; results shown here focus on the 2017 water year. Shading around the ensemble means extend to the 10th and 90th percentiles. Note that model performance in streamflow improves as the wet season progresses, and that the simulation of rock moisture is generally more dynamic than the observed time series (see potential reasoning in section 5.2). The *y*-axis of the left plot is restricted for easier viewing of observed and modelled streamflow dynamics (the wet season 2017 peak in simulated streamflow was approximately 5 mm/h)

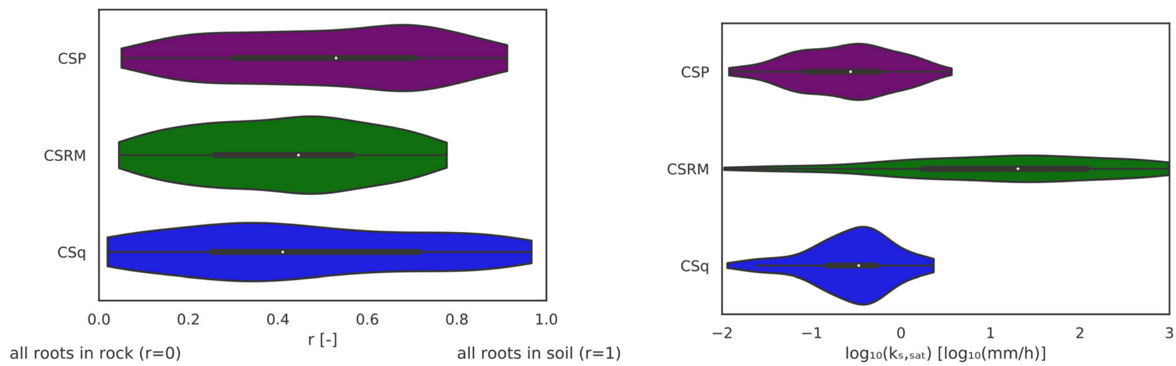

FIGURE 9 Distributions in optimal parameter values of *r* and $k_{s,sat}$, resulting from three calibration scenarios: Calibration scenario Pareto (CSP), calibration scenario rock moisture (CSRM), and calibration scenario streamflow (CSq). Each violin plot contains parameter values from the parameter sets yielding performance within the top 1% of all parameter sets for each calibration scenario. Note that values for $k_{s,sat}$ have been logarithmically spaced for clarity

periods, is 0.12 (from 0.770 to 0.653); and the in KGE mean values for rock moisture in the Pareto calibration scenario, between calibration and validation periods, actually increases by 0.02 (from 0.764 to 0.785). Due to the small changes in mean KGE between the calibration and validation periods, we can conclude that the optimal parameter sets are likely to be representative of the catchment (at least when considering the climates of the 2017 and 2018 water years).

5 DISCUSSION

5.1 Relevance of increased physical realism and restricted parameter space via Pareto optimization

The key results of our modelling study are:

- optimizing on both streamflow and dynamic rock moisture can yield parameter sets which are able to simulate both variables accurately, in terms of KGE. In contrast, calibration on streamflow alone does not reliably yield an accurate simulation of rock moisture dynamics. Therefore, calibrating on both dynamic rock moisture and streamflow yields an increase in the accuracy of modelled streamflow and rock moisture when compared to calibration on streamflow alone.
- optimal parameter values can be influenced by the choice of which variables to include in calibration, where the optimal parameter space can be restricted via calibrating on multiple targets. Specifically, Pareto optimization is shown to eliminate some physically unrealistic optima in *r*, whereas calibration on streamflow alone can yield unrealistic optima in this parameter. The parameter $k_{s,sat}$ is also highly sensitive to choice of calibration strategy.

These findings have multiple implications. First, having models that can simultaneously predict multiple hydrologically relevant variables to a good degree of accuracy has clear relevance in decision making or water management when multiple sources of water must be taken into account. Given the ecological relevance of dynamic rock moisture in semi-arid environments (Dawson et al., 2020; Hahm, Rempe, et al., 2019), earth system modellers and water managers may be interested in predicting this variable, along with streamflow. It is also clear from this study that calibration on streamflow alone is insufficient to provide accurate results in other fluxes or state variables; if a model calibrated solely on streamflow was used to make a prediction of rock moisture, it would likely be inaccurate. While our simulations of dynamic rock moisture resulting from Pareto optimization are clearly more dynamic than observations (as in Figure 8), these are at least more accurate than simulations of rock moisture obtained via calibrating on streamflow alone, as evidenced by the relatively high KGE of the rock moisture simulations calibrated on Pareto optimality. When dynamic rock moisture simulations were calibrated on dynamic rock moisture alone, the overly extreme dynamic began to attenuate (not shown), although we offer further potential avenues for accurate rock moisture modelling in the Discussion.

Second, the restriction of the parameter space (at least in some cases to physically realistic values) has promising implications for climate change studies conducted with hydrological models. While calibration in streamflow only is able to yield the highest KGE in streamflow predictions for our given modelling period, this calibration strategy yields state variables and parameters that could be unrealistic (for example, an r value of 1). Physically misrepresentative parameter sets may yield misguided projections for water availability under novel forcing regimes expected under climate change scenarios (or, conversely, as Bathurst et al. (2004) demonstrate, MCA can enable models to appropriately model the effect of a changing climate on the hydrologic system). Further, knowing the limits of the optimal parameter space can restrict the parameter search space in subsequent modelling efforts. This could lead to quicker calibration times, and depending on calibration method, more accurate model results.

5.2 Future opportunities for physically realistic or more conceptually advanced hydrologic modelling of the Elder Creek watershed

While we can claim that Pareto optimization on dynamic rock moisture and streamflow has increased the accuracy of some model results, we cannot claim that the model's overall accuracy has been increased. It is possible that while we have made *part* of the simulation more accurate, other fluxes or state variables for which we do not presently incorporate detailed observations (e.g., groundwater storage, actual evapotranspiration, or soil moisture) could have been made less accurate. Further, we would like to indicate that this model, while based on observations from Elder Creek and corresponding flow generation mechanisms, is a conceptual lumped model. Therefore it should not be considered to be a physics-based model, and could be physically (or even conceptually) more advanced.

Accordingly, there are several opportunities for improvement. First, we use the Hargreaves equation to simulate potential evapotranspiration. This equation, while easy to use, does not have the sophistication of other ET formulations, such as Priestly-Taylor or Penman-Montieth (Sumner & Jacobs, 2005), and the extent to which soil evaporation applies to the bedrock vadose zone has not been quantified. Further, the Elder Creek experimental catchment has data on sap flow in trees which in theory could be used to accurately estimate actual ET and describe spatial heterogeneity in transpirative demand (Bilir et al., 2021; Link et al., 2014). Therefore, a more advanced description of PET and plant water use is in theory possible. Second, while our approach to modelling interception has been previously used (Dralle et al., 2018), this method could be further developed as well (Allen et al., 2017; Carlyle-Moses & Gash, 2011; Junior et al., 2019; Muzylo et al., 2009). Relatedly, the vegetation cover and its impact on plant water use (Miyata et al., 2019) is not incorporated in a detailed way. Third, the focus of this model is not on hydrologic representations of soil, and accordingly the simple power law used for soils in this model is an oversimplification of the physical reality of the modelled catchment. Specifically, preferential flow paths in the soils near Elder Creek (Dietrich et al., 2013b; Langston et al., 2011) and the effect of capillarity on flow through the vadose zone are unaccounted for here. Further, the effect of slope aspect variation on various hydrologic feedbacks in the critical zone (Pelletier et al., 2018) remains unexplored. Fourth, baseflow recession can be more accurately simulated when the saturated zone is represented with a distribution of reservoir behaviours, rather than a discrete number of conventional groundwater reservoirs (Dralle et al., 2021). Fifth, while we use a power law to simulate gravity drainage from soil moisture to groundwater, and this should be a more or less plausible explanation of flow through unsaturated porous media (Clapp & Hornberger, 1978), this relatively unsophisticated method might cause some of the overly dynamic behaviour in rock moisture simulations. Alternative formulations for drainage to groundwater which specifically take into account the physics or hydrologic representation of fractured unsaturated rock can be formulated or incorporated (this topic has received prior attention; see Ebel et al. (2010); Ebel et al. (2008); Guo et al. (2019); Le Bourgeois et al. (2016); Nimmo (2007); Vrettas and Fung (2015, 2017)). In an earlier stage of the present model, we included a preferential bypass flux directly from soil to groundwater, whose magnitude was dependent on the relative rock moisture. Such a formulation can be seen as a form of semi-distribution; if one assumes that rock moisture drains to the groundwater only above a given threshold of its relative value and does so very suddenly (i.e., a large value of b_r is used), and not all of the bedrock vadose zone throughout the catchment reaches this relative moisture content at the same time, then preferential bypass effectively simulates part of the bedrock vadose zone draining to the groundwater. The limitation of this approach is that it does not allow for direct flow from the soil to the bedrock vadose zone, which is observed.

The most straightforward (and we believe, most impactful) improvement to this model in terms of increasing physical realism or accuracy would be making the model spatially distributed; this would likely have the effect of a less dynamic simulated rock moisture signal, as well as better streamflow simulations early in the wet season. The Elder Creek catchment is in many ways highly spatially heterogeneous. This is especially true with respect to the thickness of the bedrock vadose zone, which increases with distance from river channels (Rempe & Dietrich, 2014), and approaches 0 m next to Elder Creek. Relatedly, the dynamic range of rock moisture is dependent on distance to the stream. Because the thickness of the bedrock vadose zone and its dynamic moisture range are important in terms of generating flow via drainage to the groundwater, it could be the case that distribution would allow a part of the catchment with a thinner bedrock vadose zone to become active in terms of streamflow generation while the bedrock vadose zone in the rest of the catchment is still below its flow generation threshold (e.g., Dietrich et al. (2013a)). For example, and in terms of the states and fluxes in the current model, a part of the catchment (e.g., near the creek) could have its S_r state variable at or very near its (relatively small) $S_{r,max}$ value, such that significant drainage to groundwater (and therefore streamflow generation) in that particular part of the catchment could occur via f_{rg} (or even f_{og}, if the bedrock vadose zone has achieved its maximum moisture). At the same time, another part of the catchment (e.g., near the ridgetop) could have its S_r state variable very far from its (relatively large) $S_{r,max}$ value, such that there is virtually no drainage from the bedrock vadose zone to groundwater. Currently, our model does well in predicting most streamflow peaks but does poorly in simulating the streamflow at the beginning of the wet season (see Figure 8), a time for which in reality only a fraction of the bedrock vadose zone is at its flow generation threshold. While our model lumps the characteristics of the entire catchment's bedrock vadose zone into a single state variable (which therefore depicts a simplified, lumped representation of the threshold of drainage generation), we expect that a distributed model would be able to simulate these early wet season streamflow peaks more accurately than our current model (again, if for example only part of the model's spatial domain simulated enough groundwater storage for significant streamflow). We further expect that distribution might be able to improve our simulation of rock moisture, again due to allowing spatial variability in maximum rock moisture. When part of the catchment can obtain its maximum rock moisture while the rest does not, we expect that simulations of rock moisture will become less dynamic (due to the highly nonlinear power law governing drainage to groundwater only being significantly active in a fraction of the catchment at a time, rather than the entire catchment). Spatial distribution (specifically comparison of spatially distributed model results to observations) can further reduce parameter uncertainty (Werner et al., 2005).

Finally, we draw special attention to the fact that simulations of dynamic rock moisture are greater than observations for the 2 months after the last precipitation event for a given water year. For example, between 2017–05 and 2017–07 in Figure 8, right panel, modelled dynamic rock moisture begins to significantly decrease soon after the last precipitation event, whereas observations indicate that dynamic rock moisture remains near its seasonal maximum value for approximately 6 weeks after the last rainfall event. This could be potentially explained by the simulated plant water use from the bedrock vadose zone being overestimated. Having a plant water use model that specifically takes into account seasonal variation in the relative amounts of evapotranspirative demand that are satisfied from soil and the bedrock vadose zone would likely improve this section of the simulation. It is also possible that this, in turn, would improve the overly dynamic rock moisture simulation, as it is currently possible that the calibration procedure must yield large peaks in rock moisture in order to optimize rock moisture recession and thereby maximize KGE.

5.3 A note on the role of equifinality in simulating uncalibrated model components

This lumped, conceptual model, which includes 11 parameters in calibration, is subject to heavy equifinality. That is, when many parameter sets are considered, also many parameter sets will yield equally plausibility via a given performance metric (e.g., KGE in streamflow, KGE in dynamic rock moisture, or Pareto optimization on both). When many such parameter sets exist, how are we to know which to use in future simulations? Part of the value of this study is that calibration on an additional data source, namely dynamic rock moisture, can reduce equifinailty in a parameter heavy model. Calibration including this variable can evidently reduce the breadth of the post calibration parameter space, helping us to decide which parameter sets to use. Correspondingly, the concept of equifinality also explains why many (uncalibrated) components of this model may be highly inaccurate or uncertain (along with the fact that the model structure is an oversimplification of the reality of Elder Creek). For example, observations in soil moisture indicate that the seasonal maximum of moisture storage in soils in Elder Creek should be around 100 mm (Rempe & Dietrich, 2018). From Figure 5, it is clear that our model does predict this accurately; however, this is because we leverage the seasonal maximum soil moisture observations as a parameter in our model, $S_{s,max}$, removing it from calibration. When the model was run with $S_{s,max}$ included in calibration, and with a range of 1–1000 mm, the seasonal maximum of the mean of the ensemble forecast in soil moisture was roughly equal to the middle of the parameter range of $S_{s,max}$, while simulations of streamflow and rock moisture retained similar accuracies. Further, the ensemble bandwidth for soil moisture simulations is evidently large compared to its range, especially as the dry season progresses; in essence, the equifinality problem means that soil moisture can take on a large range of values and simulations of calibrated variables (streamflow and dynamic rock moisture) can still be relatively accurate. Similarly, groundwater storages are likely to be inaccurate; groundwater states themselves are not calibrated, so as long as model outputs which are compared to observations achieve a high KGE, groundwater states can take on a variety of values. When observations of a state variable or flux are not included in calibration, model results of such state variables or fluxes cannot generally be considered to be accurate (although constraining the model via restricting parameter values via observations, as we did with $S_{s,max}$, can apparently increase accuracy). The purpose of this manuscript is not to accurately simulate all of Elder Creek, but rather to demonstrate that, even with a fairly generic (albeit observation guided) model, rock moisture has value as a calibration target. This is evidenced by the simultaneous high KGEs of dynamic rock moisture and streamflow simulations, and the restriction of the post calibration parameter space as a result of Pareto optimization. In light of the utility and scarcity of rock moisture observations, more field studies on rock moisture would be of immense benefit to the hydrologic community.

5.4 On the use of KGE as an objective function for dynamic rock moisture

The KGE performance metric enjoys wide use in hydrologic sciences. Generally, it is used to quantify quality of streamflow simulations. Here it is used to assess the quality of dynamic rock moisture simulations. To our best knowledge, this has not been attempted before, and invites the question: is KGE an appropriate objective function for rock moisture? Clearly, KGEs in dynamic RM in this manuscript exceed 0.9 and span a large range of values, implying that some rock moisture simulations are much more accurate than others. Yet, it is clear that high KGE values in rock moisture are possible while the dynamic is significantly larger than observations indicate. Establishing an appropriate objective function for rock moisture time series, or perhaps exploring the accuracy of rock moisture time series beyond single number metrics, are interesting topics for future research.

6 CONCLUSION

Despite the apparent importance of unsaturated weathered bedrock as a component of the streamflow generation mechanism in many catchments, the bedrock vadose zone is usually not implemented in catchment scale models as a hydrologically distinct state variable, nor is it often used as a calibration target. To our best knowledge, unsaturated weathered bedrock water storage has not been explicitly incorporated in a catchment model and also used as a calibration target. In this manuscript, we explore the relevance of rock moisture in the context of a lumped rainfall runoff model, formulated with previously established observations and mechanisms for the water budget of the Elder Creek catchment. We achieve this by calibrating the model on streamflow individually, rock moisture individually, and on both using the concept of Pareto optimality. We find that the model is capable of accurately simulating both streamflow and rock moisture at the same time, in terms of KGE. In contrast, while calibrating on rock moisture or streamflow alone yields a relatively high KGE simulation of the calibrated variable, calibrating on a single variable is insufficient to accurately describe others. For example, calibrating on streamflow alone does not necessarily yield accurate rock moisture simulations. Therefore, Pareto optimization offers increased accuracy in the model results in streamflow and dynamic rock moisture. Further, the posterior distributions of two of the 11 calibrated parameters are sensitive to choice of calibration method. We find that calibration on streamflow alone may yield physically unrealistic optimal parameter values. Therefore, Pareto optimization on streamflow and dynamic rock moisture can offer increased model accuracy via the rejection of physically unrealistic parameter values, or at least help to reduce parameter uncertainty via constraining the calibrated parameter space. Because of the relevance of rock moisture in the water cycle, and due to the presented modelling benefits of including rock moisture as a calibration target, we advocate for further research into bedrock vadose zone process representation, more field studies to collect data on rock moisture, and further inclusion of rock moisture in earth system models.

ACKNOWLEDGEMENTS

We acknowledge support from the National Science Foundation CZP EAR-1331940 for the Eel River Critical Zone Observatory. We thank Chris Wong, Sami Cargill, Collin Bode, Peter Steel, Maryn Sanders, Wendy Baxter, and Logan Schmidt for assistance in the field. We also thank Dr. LianTze Lim and Dan MacKinnon for outstanding LaTeX support.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Here is the hydroshare containing the data, which has been published: https://www.hydroshare.org/resource/219501de7ee648e3bb0855ef9265c608/. Its DOI is 10.4211/hs.219501de7ee648e3bb0855ef9265c608. Here is the model code, which runs with 10,000 parameter sets, generates model results, and creates some of the presentation quality graphs used in this manuscript: https://colab.research.google.com/drive/1nVANYRYYAgqOZKSi1qwSvCOT0cU7yZ?usp=sharing.

ORCID

Peter T. La Follette https://orcid.org/0000-0003-1139-7421

REFERENCES

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M.. (1998). Crop evapotranspiration-guidelines for computing crop water requirements-fao irrigation and drainage paper 56. Fao, Rome, *300*(9), D05109.

Allen, S. T., Keim, R. F., Barnard, H. R., McDonnell, J. J., & Renée Brooks, J. (2017). The role of stable isotopes in understanding rainfall interception processes: A review. *Wiley Interdisciplinary Reviews: Water*, *4*(1), e1187.

Anderton, S., Latron, J., White, S., Llorens, P., Gallart, F., Salvany, C., & O'Connell, P. (2002). Internal evaluation of a physically-based distributed model using data from a mediterranean mountain catchment. *Hydrology and Earth System Sciences*, *6*(1), 67–84.

Appels, W. M., Graham, C. B., Freer, J. E., & McDonnell, J. J. (2015). Factors affecting the spatial pattern of bedrock groundwater recharge at the hillslope scale. *Hydrological Processes*, *29*(21), 4594–4610. doi:10.1002/hyp.10481

Asano, Y., Uchida, T., & Ohte, N. (2002). Residence times and flow paths of water in steep unchannelled catchments, tanakami, Japan. *Journal of Hydrology*, *261*(1–4), 173–192.

Bathurst, J., Ewen, J., Parkin, G., O'Connell, P., & Cooper, J. (2004). Validation of catchment models for predicting land-use and climate change impacts. 3. Blind validation for internal and outlet responses. *Journal of Hydrology*, *287*(1–4), 74–94.

Bilir, T. E., Fung, I., & Dawson, T. E. (2021). Slope-aspect induced climate differences influence how water is exchanged between the land and atmosphere. *Journal of Geophysical Research: Biogeosciences*, *126*(5), e2020JG006027.

Blumstock, M., Tetzlaff, D., Malcolm, I., Nuetzmann, G., & Soulsby, C. (2015). Baseflow dynamics: Multi-tracer surveys to assess variable groundwater contributions to montane streams under low flows. *Journal of Hydrology*, *527*, 1021–1033.

Botter, G., Porporato, A., Rodriguez-Iturbe, I., & Rinaldo, A. (2007). Basin-scale soil moisture dynamics and the probabilistic characterization of carrier hydrologic flows: Slow, leaching-prone components of the hydrologic response. *Water Resources Research*, *43*(2), 5043. https://doi.org/10.1029/2006wr005043

Bouaziz, L. J. E., Aalbers, E. E., Weerts, A. H., Hegnauer, M., Buiteveld, H., Lammersen, R., Stam, J., Sprokkereef, E., Savenije, H. H. G., & Hrachowitz, M. (2021). The importance of ecosystem adaptation on hydrological model predictions in response to climate change. *Hydrology and Earth System Sciences Discussions*, *2021*, 1–39. https://doi.org/10.5194/hess-2021-204

Bouaziz, L. J. E., Thirel, G., de Boer-Euser, T., Melsen, L. A., Buitink, J., Brauer, C. C., De Niel, J., Dewals, B. J., Drogue, G., Grelier, B., Melsen, L. A., Moustakas, S., Nossent, J., Pereira, F., Sprokkereef, E., Stam, J., Weerts, A. H., Willems, P., Savenije, H. H. G., & Hrachowitz, M. (2020). Behind the scenes of streamflow model performance. *Hydrology and Earth System Sciences Discussions*, *25*(2), 1–38.

Carlyle-Moses, D. E., & Gash, J. H. (2011). Rainfall interception loss by forest canopies. In *Forest hydrology and biogeochemistry* (pp. 407–423). Springer.

Clapp, R. B., & Hornberger, G. M. (1978). Empirical equations for some soil hydraulic properties. *Water Resources Research*, *14*(4), 601–604.

Clark, M. P., & Kavetski, D. (2010). Ancient numerical daemons of conceptual hydrological modeling: 1. Fidelity and efficiency of time stepping schemes. *Water Resources Research*, *46*(10), W10510.

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., & Pasteris, P. P. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology*, *28*, 2031–2064. https://doi.org/10.1002/joc.1688

Daly, C., Smith, J. I., & Olson, K. V. (2015). Mapping atmospheric moisture Climatologies across the conterminous United States. *PLoS One*, *10*, e0141140. https://doi.org/10.1371/journal.pone.0141140

Dawson, T. E., Hahm, W. J., & Crutchfield-Peters, K. (2020). Digging deeper: What the critical zone perspective adds to the study of plant ecophysiology. *New Phytologist*, *226*(3), 666–671. https://doi.org/10.1111/nph.16410

Dietrich, W. E., Rempe, D., Oshun, J., & Salve, R. (2013a). Flow paths inside a hillslope. In *AGU Fall Meeting Abstracts*, *2013*, pp. H11I–1256.

Dietrich, W. E., Rempe, D. M., Oshun, J., & Salve, R. (2013b). Hydrologic dynamics inside hillslopes. In *2013 GSA Annual Meeting in Denver*.

Dralle, D. N., Hahm, W. J., Rempe, D. M., Karst, N. J., Thompson, S. E., & Dietrich, W. E. (2018). Quantification of the seasonal hillslope water storage that does not drive streamflow. *Hydrological Processes*, *32*(13), 1978–1992.

Dralle, D. N., Harman, C. J., Deal, E., Karst, N., Rempe, D. M., & Hahm, W. J. (2021). Between-catchment variations in recession curves explained by geological diversity. In *AGU Fall Meeting* 2021.

Druhan, J. L., Fernandez, N., Wang, J., Dietrich, W. E., & Rempe, D. (2017). Seasonal shifts in the solute ion ratios of vadose zone rock moisture from the eel river critical zone observatory. *Acta Geochimica*, *36*(3), 385–388.

Ebel, B. A., Loague, K., & Borja, R. I. (2010). The impacts of hysteresis on variably saturated hydrologic response and slope failure. *Environmental Earth Sciences*, *61*(6), 1215–1225.

Ebel, B. A., Loague, K., Montgomery, D. R., & Dietrich, W. E. (2008). Physics-based continuous simulation of long-term near-surface hydrologic response for the coos bay experimental catchment. *Water Resources Research*, *44*(7), W07417.

Ernst, W. G., & McLaughlin, R. J. (2012). Mineral parageneses, regional architecture, and tectonic evolution of Franciscan metagraywackes, Cape Mendocino-Garberville-Covelo 30′ x 60′ quadrangles, Northwest California. *Tectonics*, *31*(1), TC1001. https://doi.org/10.1029/2011TC002987

Feyen, L., Vázquez, R., Christiaens, K., Sels, O., & Feyen, J. (2000). Application of a distributed physically-based hydrological model to a medium size catchment. *Hydrology and Earth System Sciences*, *4*(1), 47–63.

Freer, J., McDonnell, J. J., Beven, K. J., Peters, N. E., Burns, D. A., Hooper, R. P., Aulenbach, B., & Kendall, C. (2002). The role of bedrock topography on subsurface storm flow. *Water Resources Research*, *38*(12), 5–1–5–16. https://doi.org/10.1029/2001wr000872

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*(6). https://doi.org/10.1029/2010WR009391

Gabrielli, C., McDonnell, J., & Jarvis, W. (2012). The role of bedrock groundwater in rainfall–runoff response at hillslope and catchment scales. *Journal of Hydrology*, *450-451*, 117–133. doi:10.1016/j.jhydrol.2012.05.023

Goulden, M. L., & Bales, R. C. (2019). California forest die-off linked to multi-year deep soil drying in 2012–2015 drought. *Nature Geoscience*, *12*(8), 632–637.

Graham, R. C., Rossi, A. M., & Hubbert, K. R. (2010). Rock to regolith conversion: Producing hospitable substrates for terrestrial ecosystems. *GSA Today*, *20*(2), 4–9.

Guo, L., Lin, H., Fan, B., Nyquist, J., Toran, L., & Mount, G. J. (2019). Preferential flow through shallow fractured bedrock and a 3d fill-and-spill model of hillslope subsurface hydrology. *Journal of Hydrology*, *576*, 430–442.

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and nse performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91.

Gupta, H. V., Sorooshian, S., & Yapo, P. O. (1998). Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resources Research*, *34*(4), 751–763.

Hahm, W. J., Dralle, D. N., Rempe, D., Bryk, A., Thompson, S., Dawson, T., & Dietrich, W. E. (2019). Low subsurface water storage capacity relative to annual rainfall decouples mediterranean plant productivity and water use from rainfall variability. *Geophysical Research Letters*, *46*(12), 6544–6553.

Hahm, W. J., Rempe, D. M., Dralle, D. N., Dawson, T. E., Lovill, S. M., Bryk, A. B., Bish, D. L., Schieber, J., & Dietrich, W. E. (2019). Lithologically controlled subsurface critical zone thickness and water storage capacity determine regional plant community composition. *Water Resources Research*, *55*(4), 3028–3055.

Hale, V. C., McDonnell, J. J., Stewart, M. K., Solomon, D. K., Doolitte, J., Ice, G. G., & Pack, R. T. (2016). Effect of bedrock permeability on stream base flow mean transit time scaling relationships: 2. Process study of storage and release. *Water Resources Research*, *52*(2), 1375–1397.

Hargreaves, G. H., & Samani, Z. A. (1985). Reference crop evapotranspiration from temperature. *Applied Engineering in Agriculture*, *1*(2), 96–99.

Haria, A. H., & Shand, P. (2004). Evidence for deep sub-surface flow routing in forested upland Wales: Implications for contaminant transport and stream flow generation. *Hydrology and Earth System Sciences*, *8*(3), 334–344.

Harman, C., Sivapalan, M., & Kumar, P. (2009). Power law catchment-scale recessions arising from heterogeneous linear small-scale dynamics. *Water Resources Research*, *45*(9). https://doi.org/10.1029/2008WR007392

Hartmann, A., Wagener, T., Rimmer, A., Lange, J., Brielmann, H., & Weiler, M. (2013). Testing the realism of model structures to identify karst system processes using water quality and quantity signatures. *Water Resources Research*, *49*(6), 3345–3358.

Heppner, C. S., Loague, K., & VanderKwaak, J. E. (2007). Long-term inhm simulations of hydrologic response and sediment transport for the r-5 catchment. *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group*, *32*(9), 1273–1292.

Hrachowitz, M., & Clark, M. P. (2017). Hess opinions: The complementary merits of competing modelling philosophies in hydrology. *Hydrology and Earth System Sciences*, *21*(8), 3953–3973.

Ireson, A., Wheater, H., Butler, A., Mathias, S., Finch, J., & Cooper, J. (2006). Hydrological processes in the chalk unsaturated zone–insights from an intensive field monitoring programme. *Journal of Hydrology*, *330*(1–2), 29–43.

Jardine, P., Sanford, W., Gwo, J., Reedy, O., Hicks, D., Riggs, J., & Bailey, W. (1999). Quantifying diffusive mass transfer in fractured shale bedrock. *Water Resources Research*, *35*(7), 2015–2030.

Jayko, A. S., Blake, M. C., McLaughlin, R. J., Ohlin, H. N., Ellen, S. D., & Kelsey, H. M. (1989). *Reconnaissance geologic map of the Covelo 30- by 60-minute Quadrangle, Northern California* (Tech. Rep. No. MF - 2001). United States Geological Survey. http://ngmdb.usgs.gov/Prodesc/proddesc327.htm00011.

Junior, J. A. J., de Mello, C. R., de Mello, J. M., Scolforo, H. F., Beskow, S., & McCarter, J. (2019). Rainfall partitioning measurement and rainfall interception modelling in a tropical semi-deciduous Atlantic forest remnant. *Agricultural and Forest Meteorology*, *275*, 170–183.

Katsuyama, M., Tani, M., & Nishimoto, S. (2010). Connection between streamwater mean residence time and bedrock groundwater recharge/discharge dynamics in weathered granite catchments. *Hydrological Processes*, *24*(16), 2287–2299.

Kavetski, D., & Clark, M. P. (2010). Ancient numerical daemons of conceptual hydrological modeling: 2. Impact of time stepping schemes on model analysis and prediction. *Water Resources Research*, *46*(10). https://doi.org/10.1029/2009WR008896

Kim, H., Dietrich, W. E., Thurnhoffer, B. M., Bishop, J. K., & Fung, I. Y. (2017). Controls on solute concentration-discharge relationships revealed by simultaneous hydrochemistry observations of hillslope runoff and stream flow: The importance of critical zone structure. *Water Resources Research*, *53*(2), 1424–1443.

Kirchner, J. W. (2009). Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resources Research*, *45*(2). https://doi.org/10.1029/2008WR006912

Klos, P. Z., Goulden, M. L., Riebe, C. S., Tague, C. L., O'Geen, A. T., Flinchum, B. A., Safeeq, M., Conklin, M. H., Hart, S. C., Berhe, A. A., Hartsough, P. C., Holbrook, W. S., & Bales, R. C. (2018). Subsurface plant-accessible water in mountain ecosystems with a mediterranean climate. *Wiley Interdisciplinary Reviews: Water*, *5*(3), e1277.

Knoben, W. J., Freer, J. E., & Woods, R. A. (2019). Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrology and Earth System Sciences*, *23*(10), 4323–4331.

Koch, J., Cornelissen, T., Fang, Z., Bogena, H., Diekkrüger, B., Kollet, S., & Stisen, S. (2016). Inter-comparison of three distributed hydrological models with respect to seasonal variability of soil moisture patterns at a small forested catchment. *Journal of Hydrology*, *533*, 234–249.

Koppa, A., Gebremichael, M., & Yeh, W. W. (2019). Multivariate calibration of large scale hydrologic models: The necessity and value of a pareto optimal approach. *Advances in Water Resources*, *130*, 129–146.

La Follette, P. T., Teuling, A. J., Addor, N., Clark, M., Jansen, K., & Melsen, L. A. (2021). Numerical daemons of hydrological models are summoned by extreme precipitation. *Hydrology and Earth System Sciences Discussions*, *25*(10), 1–32.

Langston, A. L., Tucker, G. E., Anderson, R. S., & Anderson, S. P. (2011). Exploring links between vadose zone hydrology and chemical weathering in the boulder creek critical zone observatory. *Applied Geochemistry*, *26*, S70–S71.

Le Bourgeois, O., Bouvier, C., Brunet, P., & Ayral, P.-A. (2016). Inverse modeling of soil water content to estimate the hydraulic properties of a shallow soil and the associated weathered bedrock. *Journal of Hydrology*, *541*, 116–126.

Link, P., Simonin, K., Maness, H., Oshun, J., Dawson, T., & Fung, I. (2014). Species differences in the seasonality of evergreen tree transpiration in a Mediterranean climate: Analysis of multiyear, half-hourly sap flow observations. *Water Resources Research*, *50*(3), 1869–1894. https://doi.org/10.1002/2013WR014023

Lopez, P. L., Sutanudjaja, E. H., Schellekens, J., Sterk, G., & Bierkens, M. F. (2017). Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products. *Hydrology and Earth System Sciences*, *21*(6), 3125–3144.

Lovill, S., Hahm, W. J., & Dietrich, W. E. (2018). Drainage from the critical zone: Lithologic controls on the persistence and spatial extent of wetted channels during the summer dry season. *Water Resources Research*, *54*(8), 5702–5726.

Mast, A. M., & Clow, D. W. (2000). Environmental Characteristics and Water Quality of Hydrologic Benchmark Network Stations in the Western United States, 1963-95 Circular 1173-D (Circular No. 1173-D). U.S. Geological Survey.

McDonnell, J., Evaristo, J., Bladon, K., Buttle, J., Creed, I., Dymond, S., Grant, G., Iroume, A., Jackson, C. R., Jones, J. A., Maness, T., McGuire, K. J., Scott, D. F., Segura, C., Sidle, R. C., & Tague, C. (2018). Water sustainability and watershed storage. *Nature Sustainability*, *1*(8), 378–379.

Miyata, S., Gomi, T., Sidle, R. C., Hiraoka, M., Onda, Y., Yamamoto, K., & Nonoda, T. (2019). Assessing spatially distributed infiltration capacity to evaluate storm runoff in forested catchments: Implications for hydrological connectivity. *Science of the Total Environment*, *669*, 148–159.

Montgomery, D. R., Dietrich, W. E., Torres, R., Anderson, S. P., Heffner, J. T., & Loague, K. (1997). Hydrologic response of a steep, unchanneled valley to natural and applied rainfall. *Water Resources Research*, *33*(1), 91–109.

Muzylo, A., Llorens, P., Valente, F., Keizer, J., Domingo, F., & Gash, J. (2009). A review of rainfall interception modelling. *Journal of Hydrology*, *370*(1–4), 191–206.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part i—A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290.

Nimmo, J. R. (2007). Simple predictions of maximum transport rate in unsaturated soil and rock. *Water Resources Research*, *43*(5), 5372. https://doi.org/10.1029/2006wr005372

Oshun, J., Dietrich, W. E., Dawson, T. E., & Fung, I. (2016). Dynamic, structured heterogeneity of water isotopes inside hillslopes. *Water Resources Research*, *52*(1), 164–189. https://doi.org/10.1002/2015WR017485

Oshun, J., Rempe, D. M., Link, P., Simonin, K. A., Dietrich, W. E., Dawson, T. E., & Fung, I. (2012). A look deep inside the a hillslope reveals a structured heterogeneity of isotopic reservoirs and distinct water use strategies for adjacent trees, 2012, , B33A–0498. https://ui.adsabs.harvard.edu/abs/2012AGUFM.B33A0498O. Conference Name: AGU Fall Meeting Abstracts.

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the köppen-geiger climate classification, Updated world map of the Köppen-Geiger climate classification.

Pelletier, J. D., Barron-Gafford, G. A., Gutiérrez-Jurado, H., Hinckley, E.-L. S., Istanbulluoglu, E., McGuire, L. A., Niu, G.-Y., Poulos, M. J., & Rasmussen, C. (2018). Which way do you lean? Using slope aspect variations to understand critical zone processes and feedbacks. *Earth Surface Processes and Landforms*, *43*(5), 1133–1154.

Pokhrel, P., Yilmaz, K. K., & Gupta, H. V. (2012). Multiple-criteria calibration of a distributed watershed model using spatial regularization and response signatures. *Journal of Hydrology*, *418*, 49–60.

Porporato, A., Daly, E., & Rodriguez-Iturbe, I. (2004). Soil water balance and ecosystem response to climate change. *The American Naturalist*, *164*(5), 625–632.

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*(2). https://doi.org/10.1029/2003WR002805

Rakovec, O., Kumar, R., Attinger, S., & Samaniego, L. (2016). Improving the realism of hydrologic model functioning through multivariate parameter estimation. *Water Resources Research*, *52*(10), 7779–7792.

Rakovec, O., Kumar, R., Mai, J., Cuntz, M., Thober, S., Zink, M., Attinger, S., Schäfer, D., Schrön, M., & Samaniego, L. (2016). Multiscale and multivariate evaluation of water fluxes and states over european river basins. *Journal of Hydrometeorology*, *17*(1), 287–307.

Refsgaard, J. C. (1997). Parameterisation, calibration and validation of distributed hydrological models. *Journal of Hydrology*, *198*(1–4), 69–97.

Rempe, D. M., & Dietrich, W. E. (2014). A bottom-up control on fresh-bedrock topography under landscapes. *Proceedings of the National Academy of Sciences*, *111*(18), 6576–6581.

Rempe, D. M., & Dietrich, W. E. (2018). Direct observations of rock moisture, a hidden component of the hydrologic cycle. *Proceedings of the National Academy of Sciences*, *115*(11), 2664–2669.

Riebe, C. S., Hahm, W. J., & Brantley, S. L. (2017). Controls on deep critical zone architecture: A historical review and four testable hypotheses. *Earth Surface Processes and Landforms*, *42*(1), 128–156.

Rose, K., Graham, R., & Parker, D. (2003). Water source utilization by *Pinus jeffreyi* and *Arctostaphylos patula* on thin soils over bedrock. *Oecologia*, *134*(1), 46–54. https://doi.org/10.1007/s00442-002-1084-4

Rosolem, R., Gupta, H. V., Shuttleworth, W. J., Zeng, X., & De Gonçalves, L. G. G. (2012). A fully multiple-criteria implementation of the sobol' method for parameter sensitivity analysis. *Journal of Geophysical Research: Atmospheres*, *117*(D7), D07103.

Sadeghi, M., Gao, L., Ebtehaj, A., Wigneron, J.-P., Crow, W. T., Reager, J. T., & Warrick, A. W. (2020). Retrieving global surface soil moisture from grace satellite gravity data. *Journal of Hydrology*, *584*, 124717.

Safeeq, M., Bart, R. R., Pelak, N. F., Singh, C. K., Dralle, D. N., Hartsough, P., & Wagenbrenner, J. W. (2021). How realistic are water-balance closure assumptions? A demonstration from the southern sierra critical zone observatory and kings river experimental watersheds. *Hydrological Processes*, *35*(5), e14199. https://doi.org/10.1002/hyp.14199

Salve, R., Rempe, D. M., & Dietrich, W. E. (2012). Rain, rock moisture dynamics, and the rapid response of perched groundwater in weathered, fractured argillite underlying a steep hillslope. *Water Resources Research*, *48*(11), W11528. https://doi.org/10.1029/2012WR012583

Sayama, T., McDonnell, J. J., Dhakal, A., & Sullivan, K. (2011). How much water can a watershed store? *Hydrological Processes*, *25*(25), 3899–3908. doi:10.1002/hyp.8288

Schmidt, L., & Rempe, D. (2020). Quantifying dynamic water storage in unsaturated bedrock with borehole nuclear magnetic resonance. *Geophysical Research Letters*, *47*(22), e2020GL089600.

Schoups, G., Vrugt, J., Fenicia, F., & Van De Giesen, N. (2010). Corruption of accuracy and efficiency of markov chain Monte Carlo simulation by inaccurate numerical implementation of conceptual hydrologic models. *Water Resources Research*, *46*(10), W10530.

Schwinning, S. (2010). The ecohydrology of roots in rocks. *Ecohydrology*, *3*(2), 238–245. doi:10.1002/eco.134

Seidl, M., Dietrich, W. E., Schmidt, K., & de Ploey, J. (1992). The problem of channel erosion into bedrock. Functional geomorphology, 101–124.

Simonin, K. A., Link, P., Rempe, D., Miller, S., Oshun, J., Bode, C., Dietrich, W. E., Fung, I., & Dawson, T. E. (2014). Vegetation induced changes in the stable isotope composition of near surface humidity. *Ecohydrology*, *7*(3), 936–949. https://doi.org/10.1002/eco.1420

Smith, T., McNamara, J. P., Flores, A. N., Gribb, M. M., Aishlin, P., & Benner, S. G. (2011). Small soil storage capacity limits benefit of winter snowpack to upland vegetation. *Hydrological Processes*, *25*(25), 3858–3865.

Soulsby, C., Tetzlaff, D., Van den Bedem, N., Malcolm, I., Bacon, P., & Youngson, A. (2007). Inferring groundwater influences on surface water in montane catchments from hydrochemical surveys of springs and streamwaters. *Journal of Hydrology*, *333*(2–4), 199–213.

Sumner, D. M., & Jacobs, J. M. (2005). Utility of penman–monteith, priestley–Taylor, reference evapotranspiration, and pan evaporation methods to estimate pasture evapotranspiration. *Journal of Hydrology*, *308*(1–4), 81–104.

Tune, A. K., Druhan, J. L., Wang, J., Bennett, P. C., & Rempe, D. M. (2020). Carbon dioxide production in bedrock beneath soils substantially contributes to forest carbon cycling. *Journal of Geophysical Research: Biogeosciences*, *125*(12), e2020JG005795.

Vrettas, M. D., & Fung, I. Y. (2015). Toward a new parameterization of hydraulic conductivity in climate models: Simulation of rapid groundwater fluctuations in northern California. *Journal of Advances in Modeling Earth Systems*, *7*(4), 2105–2135.

Vrettas, M. D., & Fung, I. Y. (2017). Sensitivity of transpiration to subsurface properties: Exploration with a 1-d model. *Journal of Advances in Modeling Earth Systems*, *9*(2), 1030–1045.

Wanders, N., Bierkens, M. F., de Jong, S. M., de Roo, A., & Karssenberg, D. (2014). The benefits of using remotely sensed soil moisture in parameter identification of large-scale hydrological models. *Water Resources Research*, *50*(8), 6874–6891.

Werner, M., Blazkova, S., & Petr, J. (2005). Spatially distributed observations in constraining inundation modeling uncertainties. *Hydrological Processes: An International Journal*, *19*(16), 3081–3096.

Yang, D., Goodison, B. E., Ishida, S., & Benson, C. S. (1998). Adjustment of daily precipitation data at 10 climate stations in Alaska: Application of world meteorological organization intercomparison results. *Water Resources Research*, *34*(2), 241–256.

Zipper, S. C., Carah, J. K., Dillis, C., Gleeson, T., Kerr, B., Rohde, M. M., Howard, J. K., & Zimmerman, J. K. H. (2019). Cannabis and residential groundwater pumping impacts on streamflow and ecosystems in Northern California. *Environmental Research Communications*, *1*(12), 125005. https://doi.org/10.1088/2515-7620/ab534d

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: La Follette, P. T., Hahm, W. J., Rempe, D. M., Dietrich, W. E., Brauer, C. C., Weerts, A. H., & Dralle, D. N. (2022). Multicriteria analysis on rock moisture and streamflow in a rainfall-runoff model improves accuracy of model results. *Hydrological Processes*, *36*(3), e14536. https://doi.org/10.1002/hyp.14536

APPENDIX A

A.1 | A further description of numerical method

The model shown in this manuscript uses the explicit adaptive midpoint method. This method starts by calculating the flux in each state variable using the conditions at the start of the time step and makes a prediction for state variables halfway to the end of the time step (or at the midpoint). Then, the fluxes for each state variable are calculated again using the temporary midpoint state variable values, and these fluxes are averaged with the fluxes from the first half of the time step in order to make a prediction of the state variable values at the end of the time step. This method is second-order accurate, meaning that its numerical error is proportional to the time step to the second power. Therefore, as the time step decreases, this method allows for convergence to the exact solution to the system of differential equations faster than first-order methods do.

The selected numerical method is adaptive, meaning that the time step (and therefore numerical error) is automatically reduced when error tolerance criteria are not met and enlarged when error tolerance criteria are met. The midpoint method first approximates the value of the all fluxes and state variables halfway to the end of the current time step using only the conditions at the start of the time step; the values of these fluxes are the same as calculated by the (first-order) explicit Euler technique. The first-order prediction in discharge (q_{EE}) is compared to the midpoint (second-order) prediction in discharge (q_{EM}). The absolute difference between the two predictions, e, is calculated. Then, the model checks if

$$e - \tau_r \cdot q_{EM} - \tau_a < 0 \tag{A1}$$

where τ_r is the relative error tolerance (set to a value of 0.01), and τ_a is the absolute error tolerance (set to a value of 0.01 mm h^{-1}). If e is sufficiently small to pass the check, then the time step is enlarged up to a maximum value of 1 h. If e is too large, then the time step is reduced until e passes the above test. This is known as embedded error control, due to the fact that extra flux calculations are not necessary for the sake of error control (the first-order fluxes have already been calculated in the process of calculating the second-order fluxes). A similar method for error control is employed by Clark and Kavetski (2010).