Hydrogeologic influence on changes in snowmelt runoff with climate warming: Numerical experiments on a mid-elevation catchment in the Sierra Nevada, USA

S.M. Jepsen a,*, T.C. Harmon a, M.W. Meadows b,1, C.T. Hunsaker c

a University of California Merced, School of Engineering, Science & Engineering Bldg, 2, 5200 N. Lake Road, Merced, CA 95343, USA
b University of California Merced, Sierra Nevada Research Institute, Science & Engineering Bldg, 5200 N. Lake Road, Merced, CA 95343, USA
c U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station, 2081 E. Sierra Avenue, Fresno, CA 93710, USA

A R T I C L E   I N F O

Article history:
Received 6 August 2015
Received in revised form 2 November 2015
Accepted 9 December 2015
Available online 17 December 2015
This manuscript was handled by Peter K. Kitanidis, Editor-in-Chief, with the assistance of Jian Luo, Associate Editor

Keywords:
Streamflow
Runoff
Snow
Climate warming
Hydrogeology
Hydrologic property

S U M M A R Y

The role of hydrogeology in mediating long-term changes in mountain streamflow, resulting from reduced snowfall in a potentially warmer climate, is currently not well understood. We explore this by simulating changes in stream discharge and evapotranspiration from a mid-elevation, 1-km² catchment in the southern Sierra Nevada of California (USA) in response to reduced snowfall under warmer conditions, for a plausible range in subsurface hydrologic properties. Simulations are performed using a numerical watershed model, the Penn State Integrated Hydrologic Model (PHIM), constrained by observations from a meteorological station, stream gauge, and eddy covariance tower. We predict that the fraction of precipitation occurring as snowfall would decrease from approximately 47% at current conditions to 25%, 12%, and 5% for air temperature changes of +2, +4, and +6 °C. For each of these warming scenarios, changes in mean annual discharge and evapotranspiration simulated by the different plausible soil models show large ranges relative to averages, with coefficients of variation ranging from ~3 to 3 depending on warming scenario. With warming and reduced snowfall, substrates with greater storage capacity show less soil moisture limitation on evapotranspiration during the late spring and summer, resulting in greater reductions in annual stream discharge. These findings indicate that the hydrologic response of mountain catchments to atmospheric warming and reduced snowfall may substantially vary across elevations with differing soil and regolith properties, a relationship not typically accounted for in approaches relying on space-for-time substitution. An additional implication of our results is that model simulations of annual stream discharge in response to snowfall-to-rainfall transitions may be relatively uncertain for study areas where subsurface properties are not well constrained.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The timing and quantity of stream discharge from snow-influenced mountain regions, occupied by over one-sixth of the world’s population, is vulnerable to the effects of projected climate warming (Barnett et al., 2005; Nogués-Bravo et al., 2007). In California, USA, where approximately two-thirds of developed water is supplied by streamflow from the Sierra Nevada (CA DWR, 2009, pp. 3–10), climate could warm by 1.5–4.5 °C during this century (Cayan et al., 2008). Such warming in California would substantially decrease the volume of a snowpack – by anywhere from 37% to 80% (Cayan et al., 2008) – that has seasonally stored approximately two-thirds of annual precipitation (Serreze et al., 1999). This scenario now has particular relevance given the record setting drought conditions in California that began in water year 2012 (http://water.ca.gov/waterconditions/). Previous studies assessing the effect of possible atmospheric warming on changes in annual stream discharge from the Sierra Nevada have produced variable outcomes (Table 1)—with changes ranging from +3% to less than –20%. The reasons for these different predictions are unclear because of the large number of factors that vary between studies, including climate scenarios, elevations, and prediction methodologies.

Two approaches used to assess possible changes in stream discharge (hereafter “runoff”) with projected climate warming are: (i) empirical, which employ statistical relationships between
runoff and climate, and (ii) watershed modeling, which simulate the physical processes governing water and energy flows across the landscape. Predicted changes in runoff from these two approaches tend to differ in magnitude. For a 4 °C increase in air temperature and no change in precipitation, decreases in annual runoff exceeding 25% are predicted using empirical approaches, while lesser changes are generally predicted using watershed models (Table 1). Such decreases in runoff under conditions of constant precipitation would be caused by increases in evapotranspiration, which in turn would be influenced by changes in soil and/or land cover. Some of the empirically based predictions in Table 1 (i.e., Berghuijs et al., 2014; Goulden and Bales, 2014) used the space-for-time substitution approach (see Lester et al., 2014), which implicitly assumes that all influential factors in the system, such as soil and land cover, vary in phase with air temperature. However, this assumption would not likely hold true for soil development, which has time-scales on the order of thousands of years (Holbrook et al., 2014). Subsurface hydrologic properties are known to have an important effect on magnitudes of evapotranspiration and runoff, and the response of these fluxes to climate variability (Milly, 1994; Tague et al., 2008, 2014; Lundquist and Loheide, 2011; Sandhaya et al., 2014). Subsurface hydrologic properties also play an important role in the response of vegetation to climate (Kammer et al., 2013; Moyes et al., 2013; Osborn et al., 2014); for example, vegetation may be sustained through dry periods using water supplies in weathered bedrock (Witty et al., 2003; Bales et al., 2011; Kitajima et al., 2013). However, subsurface properties must not always be limiting, as demonstrated by fast and flexible shifting of vegetation across mountain landscapes in response to climate (Cannone et al., 2007; Kullman, 2008; Lenoir et al., 2008). Previous studies have contributed to the mechanistic understanding of how shifts in timing of snowmelt and/or rainfall affect changes in subsurface storage and water fluxes to streams (e.g., Milly, 1994; Risbey and Entekhabi, 1996; Tague et al., 2008; Null et al., 2010; Huntington and Niswonger, 2012). However, the extent to which subsurface hydrologic properties will mediate the effect of climate warming on long-term evapotranspiration and runoff from snow-influenced mountain regions remains unclear.

Our goal is to better understand the spatially varying effects of climate change on the hydrology and ecology of snow-influenced mountain regions. In particular, we are interested in how ecohydrologic responses to climate change may differ with hydrogeology and the partitioning of precipitation between rainfall and snowfall. Our objective here is to quantify the coupled effects of reduced snowfall (with warming) and subsurface hydrologic properties on long-term changes in runoff and evapotranspiration. Our hypothesis is that in areas where increased air temperature leads to a substantial reduction in snowfall, and advance in timing of snowmelt (earlier in the year), subsurface hydrologic properties can play a key role in long-term changes in runoff and evapotranspiration. We test this hypothesis using physically based numerical models that simulate interactions between flow-paths of water at the hillslope scale, and changes in timing and rates of surface water inputs and potential evapotranspiration. We apply our models to a small (1 km²), mid-elevation catchment in the southern Sierra Nevada (USA), Providence Creek, using catchment-averaged characteristics and meteorological forcings. Such simplifications are reasonable considering the substantial uncertainty in spatial averages of important characteristics, such as subsurface storage capacity (Bales et al., 2011; Holbrook et al., 2014) and evapotranspiration (Goulden et al., 1996), relative to their likely spatial variability at this scale. We also discuss the implications of our findings for possible large-scale differences in ecohydrologic responses across elevations with varying soil and regolith properties.
2. Study area

The study area is the 1-km² Providence Creek P301 catchment in the Sierra National Forest, California (Fig. 1a). This catchment is one of eight in the Kings River Experimental Watersheds (KREW) operated by the U.S. Department of Agriculture’s Pacific Southwest Research Station, and is part of the National Science Foundation’s Southern Sierra Critical Zone Observatory (Bales et al., 2011; Hunsaker et al., 2012). Surface elevations in the P301 catchment range from 1755 to 2114 m (mean: 1974 m, median: 1992 m), on slopes ranging from 0.1° to 39.4° (median 11.3°). The catchment is located in the “rain-snow transition zone,” exhibiting large interannual variability in the fraction of precipitation as snowfall (~35–60%, Hunsaker et al., 2012), making this location well suited for observing the effects of reduced snowfall. Climate is Mediterranean, with most precipitation occurring in frontal storms during fall to early spring. Data for model forcing and calibration were obtained from a meteorological station (Hunsaker et al., 2012) at elevation 1981 m, located approximately 1.7 km from the center of the P301 catchment, and an eddy covariance tower (Goulden et al., 2012) (hereafter “flux tower”) located near the eastern catchment boundary (Fig. 1a). Water year 2009 was the first with continuous data from the flux tower. Stream discharge at the catchment outlet was measured every 15 min at a stream gauge consisting of two nested Parshall–Montana flumes for resolving a wide range of flow rates (Hunsaker et al., 2012) (Fig. 1a). During the 2005–2011 model simulation period, 93% of annual precipitation (averaging 151 cm yr⁻¹) occurred between October 01 and April 30. During this same period, air temperatures averaged 4.1 °C between October 01 and April 30, and 15.5 °C between May 01 and September 30. Annual stream discharge averaged 54 cm yr⁻¹ during the 2005–2011 period, and annual ET averaged 76 cm yr⁻¹ during the latter three years of this period (Supplement 1, Table S1–1).

The study area is located in a Sierran mixed-conifer forest consisting of white fir, ponderosa pine, sugar pine, and incense cedar, with mature coniferous trees reaching 40–60 m in height (Hunsaker et al., 2012). On average, forest cover is approximately 16 m in height and occupies approximately 69–82% of the watershed area (Johnson et al., 2011; Harpold et al., 2014). Soils are the thinly distributed residuum and colluvium of weathered granitic rock, consisting of loamy sands and sandy loams, gravelly to cobblely (Bales et al., 2011; Johnson et al., 2011), of primarily the Gerle–Cagwin (~80%, frigid temperature regime) and Shaver (~13%, mesic temperature regime) soil families (Giger and Schmitt, 1993; USDA, 2013). Effective rooting depths range from 0.5 to 2.0 m (Giger and Schmitt, 1993). The soils transition to saprolite (paralithic contact) at a spatially averaged depth of approximately 1.0 m in the P301 catchment (USDA, 2013). The saprolite and deeper weathered bedrock provide an important water reservoir for summer and fall evapotranspiration and baseflow (Witty et al., 2003; Bales et al., 2011). The subsurface storage capacity of saprolite below a transect in the northern P301 catchment was found to be 3 m³ m⁻² (Holbrook et al., 2014), an order of magnitude greater than that of the soil located above the paralithic contact.

3. Methods

3.1. Model overview

Stream discharge (runoff) from the catchment was simulated for water years 2005–2011 using the Penn State Integrated Hydrologic Model (PIHM) version 2.2, which uses a semidiscrete finite volume method to couple the evolution of hydrologic states on an unstructured numerical mesh (Qu and Duffy, 2007; Kumar, 2009). Lateral groundwater flow in PIHM is computed using Darcy’s Law with the Dupuit–Forchheimer assumption (no vertical flow). Fluxes of infiltration and groundwater recharge are computed using analytic solutions derived from van Genuchten relationships (van Genuchten, 1980) and hydraulic head differences between surface water, vadose zone, and phreatic zone (Kumar, 2009). This method accounts for the generation of both infiltration excess and saturation excess overland flow. Computations of evapotranspiration followed from the Penman and Penman–Monteith equations, with functional terms depending on weather (air temperature, vapor pressure deficit), soil moisture, interception storage, and vegetation properties (Kumar, 2009). The mesh of our model consisted of 256 prismatic elements and 16 linear stream elements (Fig. 1b), and was generated using an open-source GIS interface (PIHMgis) that accompanied PIHM version 2.2 (Kumar et al., 2009).

We simulated subsurface hydrologic properties using different models of spatially uniform soil thickness and hydrologic properties. These “soils” actually represent a composite of soil, saprolite, and weathered bedrock. We modeled soil thicknesses of 1.5, 3.0, 4.5, and 6.0 m to encompass a plausible range in subsurface storage capacities for the study area (Bales et al., 2011; Holbrook et al., 2014). We forced models with hourly air temperature, relative humidity, wind speed, downwelling solar radiation, and precipitation (P) from the meteorological station (Fig. 1a). The precipitation values were uniformly disaggregated from daily measured totals. Three atmospheric warming scenarios were constructed by uniformly shifting measured 2005–2011 air temperatures upward by 2, 4, and 6 °C, which encompasses a possible range in long-term warming scenarios based on GCM predictions (Cayan et al., 2008). Precipitation was maintained at 2005–2011 measured values (i.e., not changed), pursuant to our objective of singling out the effects of warming and also reasonable given the broad range of GCM predictions (Ficklin et al., 2012). The modeled atmospheric warming scenarios were initialized with subsurface storages from a four-year spin up to steady-state conditions for water year 2009 (calibration year), with modified air temperature incorporated into the spin-up forcings. Model time steps were one minute for convergence of solutions (via the SUNDIALS package), and hourly for model output.

3.2. Model calibration

We manually calibrated models by sampling all combinations of parameters being estimated, using a Linux cluster of 15 processors (multi-processing in Python). Considering this method, and that run times were approximately 0.5 h (model year)⁻¹, we decided to calibrate models to a single water year (2009) and estimate no more than three model parameters at a time (<10 bins per parameter). We divided the calibration process into two steps, summarized here and further detailed in Supplement 2. In the first step (Step 1), adopting the approach of Kumar et al. (2013), we calibrated horizontal groundwater flow to hydrograph recession by optimizing three parameters controlling the variation of transmissivity with depth: horizontal hydraulic conductivity of soil matrix and macropores, and depth of macropores. This was done by minimizing the RMS model error in logarithm of stream discharge (Q) during July–September streamflow recession, taking as measured quantities the time-integrated relationship between observed values of dQ/dt and Q. In the second step (Step 2, detailed below), we calibrated model infiltration by identifying and then estimating three parameters having the greatest influence on modeled stream discharge over the entire 2009 water year. To identify these influential parameters, we conducted a global sensitivity analysis using the Method of Sobol’ (Saltelli et al., 2010), implemented in the SALib sensitivity analysis library for Python (Herman, 2014). This
sensitivity analysis consisted of 1000 total runs sampling nine different model parameters (Supplement 2, Table S2-4). Based on this analysis, the parameters most influencing annual discharge and RMS error in discharge were the van Genuchten water retention parameters \( \text{Alpha}, \text{Beta}, \text{and ThetaE} \) (difference between total porosity and residual porosity), with total effects accounting for over 98% of model variance. These parameters, also identified by Shi et al. (2013) to be most influential, do not vary with depth in PIHM version 2.2.

In each model run of calibration Step 2 (infiltration), we applied an ET-correction loop based on Newton’s method to spin up models to water year 2009 and bring annual modeled ET to the observed (flux tower) value of 75 cm (Supplement 2). For the objective function of calibration, we initially tried using the Nash–Sutcliffe efficiency (NS) of daily modeled discharge, but were unable to obtain acceptable model fits (Table 2). We attributed this to a positive bias in modeled net flow of water to the catchment (i.e., influxes minus outfluxes), indicated by large observed “residuals” in \( P – Q – \text{ET} \). During water years 2009–2012, the average observed value of \( P – ET, 73 \text{ cm yr}^{-1} \), was 43% greater than the average observed value of discharge (Q), 51 cm yr\(^{-1}\), an imbalance not likely caused by increases in storage (Supplement 1). We were unable to capture this degree of imbalance between \( P – \text{ET} \) (constrained) and Q (simulated) in our Step 2 calibrations. We attributed this problem to two possible causes: (i) observations of \( P, Q \), and ET not representing all water fluxes to and from the P301 catchment with sufficient accuracy relative to magnitudes of discharge, and (ii) the difficulty of representing catchment water balance in single-year model calibrations. Our solution to this water balance problem was to calibrate models to the shape of the hydrograph by maximizing the Nash–Sutcliffe efficiency of daily modeled discharge for both modeled and observed hydrographs normalized to unity (\( NS_{\text{norm}} \)). We later validated all models over the 2005–2011 time period using full (unnormalized) discharge values at the daily, monthly, and yearly time-scale, using the following metrics: Nash–Sutcliffe efficiency (NS), coefficient of determination (\( R^2 \)), percent bias (BIAS), and fraction of modeled discharge as baseflow (BSFLW\(_{\text{mod}}\) (Table 3). We also examined error statistics of annual modeled discharge offset by a “bias-correction” of −22 cm to account for the observed imbalance between \( P – \text{ET} \) and Q averaged over water years 2009–2012 (see Supplement 1). The NS and BIAS metrics for these annual bias-corrected values are denoted as \( NS_{\text{corr}} \) and \( \text{BIAS}_{\text{corr}} \), respectively (Table 3). Model 2005–2011 validation runs were initialized with subsurface water storages from a four-year spin up to steady-state conditions for the calibration year (2009), followed by relaxation spin-up to the initial observed discharge in water year 2005.

Uncertainty in model predictions of annual runoff and ET in association with the water retention parameters, sampled in calibration Step 2 (see above), were examined using the Generalized Likelihood Uncertainty Estimation (GLUE) method (Beven and Binley, 1992). We examined these uncertainties in order to compare the range of predictions associated with different water retention parameters to the range of predictions for the different soil thicknesses. The probability of a runoff or ET prediction using set \( i \) of the water retention parameters was assumed to be proportional to the objective function of calibration Step 2, \( (NS_{\text{norm}})_i \) (see above):

\[
p_i = (NS_{\text{norm}})_i / \sum_{j=1}^{S} (NS_{\text{norm}})_j, \tag{1}
\]

![Fig. 1. Location of the Providence P301 study catchment in the Sierra National Forest, California, shown alongside the mesh of the watershed model. (a) Topographic map showing the catchment boundary and locations of instruments used to calibrate and force models (DEM from Guo et al., 2010), (b) mesh of the watershed model generated using PIHMgfs, consisting of 256 prismatic elements and 16 linear stream elements.](image-url)
Table 3
Validation error statistics of modeled annual, monthly, and daily stream discharge over the 2005–2011 period. Upper panel lists Nash-Sutcliffe efficiencies (NS) and $R^2$ values, lower panel lists statistics relating to model bias and baseflow. NScor = Nash-Sutcliffe efficiency of modeled annual discharge offset by –22 cm bias correction, BIAScorr = percent error in modeled annual discharge offset by –22 cm bias correction, BSFLWmod = fraction of modeled discharge as base flow; BSFLWobs = fraction of observed discharge as base flow, estimated using digital filter in Arnold et al. (1995).

<table>
<thead>
<tr>
<th>$t_{500}$ (m)</th>
<th>NS</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yearly</td>
<td>Monthly</td>
</tr>
<tr>
<td>1.5</td>
<td>0.82</td>
<td>0.71</td>
</tr>
<tr>
<td>3.0</td>
<td>0.83</td>
<td>0.85</td>
</tr>
<tr>
<td>4.5</td>
<td>0.78</td>
<td>0.80</td>
</tr>
<tr>
<td>6.0</td>
<td>0.73</td>
<td>0.74</td>
</tr>
</tbody>
</table>

For each modeled soil thickness and air temperature scenario, five different water retention parameter sets were sampled ($i = 1...5$) corresponding to the 20th, 40th, 60th, 80th, and 100th percentiles of the cumulative likelihood distribution of the Step 2 calibration (see Supplement 2, Table S2-2 for parameter sets). This required a total of 80 forward model runs over the 2005–2011 time period. The mean, $\mu$, and standard deviation, $\sigma$, of a model prediction of runoff or ET were calculated as:

$$\mu = \frac{\sum_i p_i x_i}{\sum_i p_i},$$

$$\sigma = \sqrt{\sum_i p_i (x_i - \mu)^2},$$

where $x_i$ is a prediction based on model parameter set $i$.

3.3. Snow model

PHM version 2.2 uses the temperature-index method to simulate the accumulation and melt of a snowpack on each element of the mesh. The spatial variability of a snowpack could in general be simulated by providing zones of meteorological observations in the PHM forcing file (however, single zone used in this study). Hourly snowmelt was computed as the product of a melt factor and the air temperature departure above a threshold value for melt (Supplement 3) (Kumar, 2009). New snowfall was simulated based on measured precipitation and air temperature relative to threshold values for all snowfall (Ts) and all rainfall (Tr) (Supplement 3).

We calibrated the snow model to the snow pillow, located at the meteorological station, for each water year of simulation (Supplement 3). During these calibrations, it became apparent that the snow pillow substantially overestimated basin-average snow water equivalent (SWE), by approximately 54% in water year 2010 based on LiDAR mapping (Harpod et al., 2014) and 82% in water year 2011 based on a snow survey (41 observations) (Supplement 3, Table S3-2). To account for this bias in our snow model, we incorporated a scaling factor (Tr/S) into the PHM calibrations that additively shifted the snowfall–rainfall parameters Ts and Tr. This had the effect of scaling the SWE hydrographs (Supplement 3, Fig. S3-1). Using snowmelt models with best-fit Tr/S-values, our prediction errors in basin-average maximum SWE were +14% in water year 2010 and +21% in water year 2011, less than one-third of the over-representation of basin-average SWE at the snow pillow (Supplement 3, Table S3-2). These snowmelt models also provided reasonable predictions of reduced SWE with increased air temperature. For the 4 °C warming case, our models predicted an 83% reduction in maximum SWE (Supplement 3, Table S3-3), compared to a similar prediction of 79–93% by Cayan et al. (2008) for similar elevations in the Sierra Nevada.

4. Results

4.1. Model assessment

The calibrated models of all soil thicknesses (1.5–6.0 m) adequately explained the observed variability in yearly and monthly stream flows (NS > 0.5; Moriasi et al., 2007) (Table 3). At the daily time-scale, the model with the 3.0 m soil thickness performed best (NS = 0.69) and the model with the 1.5 m soil thickness performed worst (NS = 0.31) (Table 3). The model with the 1.5 m soil thickness also predicted insufficient baseflows and overly flashy discharge relative to observations (Fig. 2, Table 3). All calibrated models underestimated the relatively low stream discharge values that occur at probabilities of exceedance greater than 80% (Fig. 3), particularly the thinnest (1.5 m) soil model.

We also compared modeled and observed trends in annual P–Q (P = precipitation, Q = discharge) versus annual air temperature and annual precipitation, under the rationale that P–Q constitutes a different portion of the catchment water balance (i.e., P–Q = ET + storage change) than Q alone, and also because annual P–Q is well correlated to annual ET ($R^2 = 0.78$, p-value = 0.001 [two-tailed], not shown) in the surrounding upper Kings River watershed (Goulden et al., 2012). The observed trend in P–Q versus annual precipitation was well captured by the 6.0 m soil thickness model (Fig. 4, bottom left). The observed trend in P–Q versus annual air temperature was well captured by the 4.5 m soil thickness model (Fig. 4, second-to-bottom right). This finding suggests that the thicker soil models are better capturing the water-balance response of the catchment to climate variability than the thinner soil models. All calibrated models showed large percent biases in discharge, ranging from +22% to +33% (Table 3), which we attribute at least partly to a positive error bias in observed net fluxes of water to the catchment (see Section 3.2). With the “bias correction” of –22 cm applied to annual modeled discharge to account for the large observed residuals in P–Q–ET during 2009–2012 (see Section 3.2), the thinnest soil model produced the highest Nash–Sutcliffe efficiency of annual flow (NScor) and the lowest percent bias (BIAScor) (Table 3). This further suggests that the thicker soil models are best capturing the water-balance response of the catchment to climate variability.

4.2. Sensitivity of changes in annual runoff to subsurface hydrologic properties

Simulated changes in mean annual runoff and ET with increased air temperature depend strongly on subsurface hydrologic properties. The thinnest soil models (1.5 m) all produced increased runoff, with a maximum change of +5% (relative to 54 cm yr$^{-1}$ baseline) for the intermediate warming scenario (Fig. 5a). In contrast, the thicker soil models generally produced decreased runoff, with a maximum decrease of –14% for the thickest soil and greatest warming scenario (+6 °C) (Fig. 5a). These simulated changes in runoff were driven by changes in ET (Fig. 5b), which are opposite in sign and equal in magnitude to changes in runoff (Fig. 5a versus b). The ranges in annual runoff and ET from the uncertainty analysis, which sampled the range in water retention parameters (Section 3.2, calibration Step 2), are shown by the vertical error bars in Fig. 5a and b. These ranges span up to nearly half of the total range in runoff and ET predictions for the different soil thicknesses (Fig. 5a and b), showing that both subsurface storage capacity and
unsaturated hydraulic properties can play an important role in long-term changes in runoff and ET with reduced snowfall. Coefficients of variation in predicted changes in annual runoff, evaluated across the different soil thicknesses, were 3, –3, –1 for air temperature increases of 2, 4, and 6 °C (Fig. 5a), demonstrating that the effect of subsurface hydrologic properties on annual runoff can be greater than the effect of increased air temperature alone.

What is driving such large variation in annual runoff across the different soil models? Warming of the atmosphere would result in less precipitation falling as snow (47–89% reduction) and earlier melting of snowpacks (~2–4 weeks) (Supplement 3, Table S3-3). This, in turn, can induce soil moisture deficits during the growing season that limit increases in annual ET and therefore buffer decreases in runoff (Jeton et al., 1996; Dettinger et al., 2004; Tague et al., 2009). In our model results, the greatest soil moisture limitations on ET occurred in the thinnest soils during early May (Fig. 5c), when decreases in ET reached nearly 0.3 cm day⁻¹. Lesser limitations on ET were simulated in the thicker soils (Fig. 5c), because these have a greater capacity to store inputs of early-season surface water. These differing ET responses are strongly coupled to reductions in snowfall and advances in snowmelt. Simulations applying increases in air temperature, but maintaining all snowfall and snowmelt dynamics at their baseline 2005–2011 values, predicted runoff changes that are relatively insensitive to soil thickness, with coefficients of variation not exceeding 0.1 when evaluated across the different soil thicknesses (Fig. 5d). Therefore, subsurface hydrologic properties play a particularly important role in long-term runoff and ET when climate warming brings about a substantial transition from snowfall to rainfall.

4.3. Sensitivity of changes in runoff timing to subsurface hydrologic properties

Our simulations showed advances in timing of runoff by approximately 20, 29, and 31 days for air temperature increases of 2, 4, and 6 °C, respectively (Fig. 6). Changes were similar (within
runoff evaluated across the different soil thicknesses ranged from –0.03 to –0.1 (Fig. 6), an order of magnitude lower than the corresponding variability in annual runoff and ET (CV ranging from –3 to 3, Section 4.2). Therefore, changes in runoff timing with reduced snowfall are much less sensitive to soil thickness than are changes in runoff quantity (presented in Section 4.2).

Modeled timing of early season runoff is substantially more sensitive to subsurface hydrologic properties than mid- to late-season runoff. For the intermediate air warming case of +4°C, a two-week range occurred in the modeled timing of 25% cumulative runoff in association with the different soil thicknesses and water retention parameters (Fig. 6a). The timing of this 25% cumulative runoff occurred between February 25 and March 17 at baseline 2006–2011 conditions (not shown). Others have drawn attention to the elevated flood hazards this time of the year in association with heavy rainfall events (Lettenmaier and Gan, 1990; Dettinger et al., 2004).

5. Discussion

5.1. Elevational dependence of runoff response to climate warming

Our findings suggest that the effects of atmospheric warming on long-term runoff and evapotranspiration in snow-influenced catchments of the Sierra Nevada may depend strongly on elevation. Modeled subsurface storage capacities of 1.2–2.4 m$^3$ m$^{-2}$ (Supplement 2, Table S2-2) are needed to adequately capture the

![Flow duration curves of stream discharge from observations and models (water years 2005–2011). All models underestimate stream discharge for probabilities of exceedance greater than approximately 80%, with greatest underestimates from the thinnest (1.5 m) soil model.](image)

![Difference between annual precipitation and discharge (P–Q) versus (left) annual precipitation and (right) annual air temperature. (a) Results from observations, (b) results from models (soil thicknesses labeled). Dashed lines show fits from linear regression. Trend = regression slope, p-value = statistical significance of regression (Student’s two-tailed t-distribution).](image)
observed dynamics of streamflow from the mid-elevation catchment under study (see Section 4.1). With increasing elevation from the P301 catchment, subsurface storage capacities decline markedly. At elevations above 2500 m, which occur in 50% of the upper Kings River watershed and 40% of the neighboring upper San Joaquin River watershed, the top of lithic bedrock is within approximately 25 cm of the ground surface (Fig. 7). These high elevation regions presently exhibit relatively high runoff ratios (Q/P) under the combined effects of abundant precipitation (orographic) and low ET (Hunsaker et al., 2012; Goudlen and Bales, 2014), the latter related to the presently cool air temperatures, limited vegetation cover, and rapid runoff. A study by Goudlen and Bales (2014) integrating ET observations, remote sensing products, and space-for-time substitution found that future climate warming at high
elevations of the upper Kings River watershed may bring about substantial increases in ET (∼3 cm °C⁻¹ basin-wide). This could result from expansion and/or shifts in vegetation cover and reductions in cold limitations on gross primary production, according to their study. Based on our results, this ET response to warming may be muted to some extent by the combined effects of reduced snowfall and low subsurface storage capacities at high elevation, as Fig. 5 illustrates for subsurface storage capacities ranging from 0.5 to 2.4 m³ m⁻² and snowfall ranging from 5% to 47% of annual precipitation. Though we have not accounted for possible changes in vegetation cover with climate warming, the finding that runoff and ET trajectories depend largely on subsurface hydrologic properties (Fig. 5a and b) suggests that putative expansion and/or shifts in vegetation, and reductions in cold limitation, may be largely mediated by the prevailing hydrogeologic conditions. For this argument to hold, large reductions in snowfall would be needed, based on our findings (Fig. 5a versus d). Substantial decreases in snowfall are possible with future warming at high elevations of the Sierra Nevada (Tague et al., 2009; Ficklin et al., 2012, 2013; Costa-Cabral et al., 2013), though these areas will be more resilient than lower, warmer elevations in the rain-snow transition zone (Stewart, 2009).

5.2. Considerations regarding space-for-time substitution

The space-for-time substitution approach, which has been used to predict substantial reductions in runoff with possible climate warming (Berghuijs et al., 2014; Goulden and Bales, 2014), uses empirical relationships that integrate observations across a broad range of land cover types, including vegetation and soils. In doing so, the method does not typically quantify the response times of environmental processes, such as vegetation growth/expansion and soil formation (regolith weathering), that may follow changes in climate. In regions where climate warming causes substantial reductions in snowfall (and concurrent advances in timing of snowmelt), the capacity of soils to store water for use by vegetation during the growing season can have as much effect on changes in long-term runoff and evapotranspiration as changes in air temperature alone, as our simulations of a mid-elevation catchment suggest (see Section 4.2). These results are generally consistent with observational studies of subalpine regions (Moyses et al., 2013; Osborn et al., 2014) and alpine regions (Kammer et al., 2013) showing that the water capacity of soil can play a key role in the response of vegetation to climate variability. A relevant question concerning the ultimate response of mountain hydrology to climate change is: To what extent can soil development and regolith weathering “keep up” with climate change in a way that is implicitly represented in the spatial datasets used in space-for-time substitution approaches? Time-scales of soil development are likely to be exceedingly slow, on the order of thousands of years (Holbrook et al., 2014), relative to time-scales of future climate projections. However, vegetation has been shown to perhaps respond more quickly to climate change (Cannone et al., 2007), with observed migration rates ranging from tens of meters (Kelly and Goulden, 2008; Lenoir et al., 2008) to hundreds of meters (Kullman, 2008) elevation per decade. These migration rates may depend to some extent on pre-existing subsurface hydrologic properties. Increased information about subsurface properties in studies documenting the expansion of mountain vegetation with climate change would provide useful information for further examining these relationships.

5.3. Subsurface hydrologic properties as a potential source of model uncertainty

Two findings together suggest that subsurface hydrologic properties may be an important source of uncertainty in model predictions of long-term runoff and ET in snow-influenced catchments: (i) model predictions are highly sensitive to subsurface hydrologic properties whenever increases in air temperature lead to a substantial reduction in snowfall (Section 4.2); (ii) subsurface properties tend to be relatively unconstrained in mountainous areas underlain by thin soils and thick regolith (saprolite and weathered bedrock) (Witty et al., 2003; Holbrook et al., 2014). Milly (1994) had similar findings regarding the importance of subsurface properties, showing that runoff prediction uncertainties associated with subsurface water holding capacity can be nearly as great as uncertainties associated with precipitation and
potential evapotranspiration (contributions to model error variance from water holding capacity, precipitation, and potential evapotranspiration equaling 25%, 31%, and 31%, respectively). This source of uncertainty may be contributing to the notable range, approximately 25% relative to baseline values, in model predictions of annual runoff from snow-influenced catchments under air temperature increase scenarios (Lettenmaier and Gan, 1990; Jeton et al., 1996; Tague et al., 2009; Null et al., 2010; He et al., 2013; Tague and Peng, 2013) (Table 1).

6. Conclusions

Long-term changes in annual runoff and evapotranspiration (ET) in response to climate warming are found to be highly sensitive to subsurface hydrologic properties when significant snowfall to rainfall transitions occur. We estimate that a 2–6 °C increase in air temperature over a 1-km², mid-elevation catchment in the southern Sierra Nevada would result in a 47–89% reduction in snowfall. For such warming and shift in phase of precipitation, simulated trajectories of annual runoff and ET vary greatly over a plausible range in subsurface hydrologic properties for the study area. For example, coefficients of variation in simulated annual runoff change, evaluated across a plausible range in subsurface storage capacities, range from 1 to 3 in magnitude, depending on climate warming scenario. Simulated changes in runoff timing are much less sensitive to subsurface hydrologic properties than are changes in runoff quantity (order of magnitude lower coefficients of variation). These results imply potentially contrasting hydrologic responses of long-term runoff and ET across elevational gradients where changes in hydrogeology, such as subsurface storage capacity, occur. For example, at the higher elevations of the upper Kings River watershed (USA) where exposed bedrock is common, a given reduction in the fraction of precipitation as snowfall (caused by atmospheric warming) could elicit opposite changes in annual runoff than at lower elevations where subsurface storage capacity is typically greater. These results should be considered in space-for-time substitution approaches that use empirical relationships between climate and runoff (and/or evapotranspiration) based on spatial datasets encompassing a range of hydrogeological conditions. Our results also suggest that in snow-influenced regions, unconstrained subsurface properties of saprolite and weathered bedrock may be an important source of model uncertainty in assessments of long-term water balance in a changing climate.

Acknowledgments

We thank the anonymous reviewers for their helpful comments on the manuscript. Technical assistance was provided by Otto Alvarez, Jacob Flanagan, Michael Goulden, Jon Herman, James Kirchner, Phil Saksa, Yuning Shi, and Xuan Yu. Funding was provided by US National Science Foundation (NSF) Award # CBE-T 1204841 (Water Sustainability & Climate Program) and by the Inter-American Institute for Global Change Research (IAI) CRN 3038 (under NSF Award # GEO-1128040). Data were provided by the Pacific Southwest Research Station, USDA Forest Service, which operates the Kings River Experimental Watersheds, and by the US NSF Southern Sierra Critical Zone Observatory.

Appendix A. Supplemental material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jhydrol.2015.12.010.