Hydrology and Earth System Sciences

Open Access

On the contribution of groundwater storage to interannual streamflow anomalies in the Colorado River basin

E. A. Rosenberg[1,2]**, E. A. Clark**[1]**, A. C. Steinemann**[1,3,4]**, and D. P. Lettenmaier**[1]

[1]Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington 98195, USA
[2]Hazen and Sawyer, P.C., New York, New York 10018, USA
[3]Evans School of Public Affairs, University of Washington, Seattle, Washington 98195, USA
[4]Scripps Institution of Oceanography, La Jolla, California 92093, USA

Correspondence to: E. A. Rosenberg (erosenberg@hazenandsawyer.com)

Abstract. We assess the significance of groundwater storage for seasonal streamflow forecasts by evaluating its contribution to interannual streamflow anomalies in the 29 tributary sub-basins of the Colorado River. Monthly and annual changes in total basin storage are simulated by two implementations of the Variable Infiltration Capacity (VIC) macroscale hydrology model – the standard release of the model, and an alternate version that has been modified to include the SIMple Groundwater Model (SIMGM), which represents an unconfined aquifer underlying the soil column. These estimates are compared to those resulting from basin-scale water balances derived exclusively from observational data and changes in terrestrial water storage from the Gravity Recovery and Climate Experiment (GRACE) satellites. Changes in simulated groundwater storage are then compared to those derived via baseflow recession analysis for 72 reference-quality watersheds. Finally, estimates are statistically analyzed for relationships to interannual streamflow anomalies, and predictive capacities are compared across storage terms. We find that both model simulations result in similar estimates of total basin storage change, that these estimates compare favorably with those obtained from basin-scale water balances and GRACE data, and that baseflow recession analyses are consistent with simulated changes in groundwater storage. Statistical analyses reveal essentially no relationship between groundwater storage and interannual streamflow anomalies, suggesting that operational seasonal streamflow forecasts, which do not account for groundwater conditions implicitly or explicitly, are likely not detrimentally affected by this omission in the Colorado River basin.

1 Introduction

Among the most important contributors to the skill of a streamflow forecast are the estimation of initial hydrologic conditions (IHCs, i.e., basin water storage at the time of forecast) and prediction of future meteorological anomalies (Mahanama et al., 2012). Despite some demonstrated skill in seasonal climate forecasting (see, e.g., Stern and Easterling, 1999; Troccoli et al., 2008), most meteorological forecasts for leads longer than about a month are of limited skill (Shukla and Lettenmaier, 2011). Thus, at seasonal lead times, accurate streamflow forecasts are possible mostly in situations where future runoff is more strongly related to catchment water storage than to meteorological anomalies during the forecast period (Wood and Lettenmaier, 2008; Shukla and Lettenmaier, 2011). In the American West, this condition is the basis for the statistical seasonal streamflow forecasts issued by the Natural Resources Conservation Service (NRCS; Garen, 1992) and is an implicit attribute of the dynamically generated ensemble streamflow predictions issued by the National Weather Service (NWS; Day, 1985).

For many rivers in the Western US, more than half of the annual streamflow is derived from snowmelt, and snow water storage has historically been the single most significant predictor for statistical streamflow forecasts (Church, 1935). The opportunity to exploit the relationship between soil moisture and runoff in statistical streamflow forecasts was also recognized in early studies (e.g., Boardman, 1936; Clyde, 1940), although accumulated precipitation is more typically used as a proxy due to a paucity of in situ soil moisture observations (Speers et al., 1996; Koster et al.,

2010). Nevertheless, recent modeling studies have demonstrated that early-season soil moisture can have a significant influence on seasonal streamflow, even where annual hydrographs are dominated by spring snowmelt (Koster et al., 2010; Mahanama et al., 2012).

One contribution to basin storage that traditionally has been neglected in streamflow forecasts is groundwater. By sustaining baseflow and laterally redistributing subterranean water, groundwater discharge provides an important link in the hydrologic cycle. With the exception of arid climates where it can be essentially disconnected from the land surface, groundwater also receives surplus during wet periods and offsets deficits during drought (Fan et al., 2007), providing the ability to carry over storage from one year to the next. Although known to be the largest of the storage terms in quantity, however, the magnitude of groundwater's temporal variability relative to that of soil moisture and snow is often poorly understood. How significant are interseasonal and interannual groundwater anomalies for seasonal streamflow forecasts?

The answer to this question has several important implications. One concerns the accuracy of operational seasonal streamflow forecasts, which do not account for groundwater conditions implicitly or explicitly. Another involves the representation of the subsurface in land surface models (LSMs). Notwithstanding their physical basis, LSMs are fundamentally simplifications of natural processes. Until recently, most have lacked a groundwater representation entirely, typically formulating lower boundary conditions either as zero flux or as drainage under gravity (Maxwell and Miller, 2005). Yet such simplifications can significantly bias the estimation of soil water flux and streamflow, and without an explicit representation of the water table, the land surface water budget may not close other than on very long time averages (Yeh and Eltahir, 2005). Consequently, a number of groundwater parameterizations have been proposed (e.g., Liang et al., 2003; Niu et al., 2007), and some research has suggested that inclusion of an explicit aquifer model can reduce the sensitivity of model performance to incorrect parameter choices (Gulden et al., 2007). Fewer studies, however, have been devoted to understanding the effect of these modifications on variations in total basin storage, particularly at the interannual scale.

The objective of this paper is to evaluate the contribution of groundwater storage to interannual streamflow anomalies, and hence to seasonal and interannual streamflow predictability, in the Colorado River basin, which is iconic in the American West. It has been described as the most regulated and over-allocated river in the world (NRC, 2007), with some recent research suggesting that current water deliveries are not sustainable (Barnett and Pierce, 2009). Yet modeling studies of the potential effects of climate change on Colorado River streamflow have been notably incongruent (Hoerling et al., 2009), which may in part suggest a misunderstanding of catchment processes, including the role of groundwater storage in the basin's hydrologic response to climatic variations and change. Thus, while our study is motivated by seasonal streamflow forecasts, an additional interest is in evaluating hydrologic models, which are typically validated only by streamflow, by providing an observation-based assessment of total basin storage anomalies.

We give particular attention to results over the last decade for several reasons. First, conditions have been especially dry in the Colorado River basin during this period (Quinlan, 2010), rendering accurate water supply forecasts particularly important. Second, focusing on the recent past permits better assessment of results from institutional memory. Finally, we are able to supplement hydroclimatic data sets over the last decade with remote sensing observations that were previously unavailable, specifically, estimates of evapotranspiration derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) and total water storage estimates based on the Gravity Recovery and Climate Experiment (GRACE) satellites.

2 Study area

The Colorado River flows for 2300 km (1450 mi) through seven US and two Mexican states to its mouth at the Gulf of California (Fig. 1). Its 630 000 km^2 (240 000 mi^2) drainage area was divided for purposes of the Colorado River Compact of 1922 (and consequently for many water management purposes) into an Upper Basin and a Lower Basin at Lee's Ferry, Arizona. The hydrograph of the Colorado is dominated by snowmelt, with roughly 70 % of its annual streamflow derived from this source. Furthermore, an estimated 85 % of its streamflow originates from just 15 % of the basin area located in the headwaters of the Southern and Middle Rocky Mountains (Christensen and Lettenmaier, 2007). The majority of the basin is comprised of desert or semiarid rangeland, which generally receives less than 250 mm (10 in.) of precipitation per year. Most precipitation in the high elevation streamflow source areas occurs in winter and spring and comes from eastward-tracking Pacific storm systems (Robson and Banta, 1995). The Colorado has a combined reservoir storage capacity of 74.0 billion cubic meters (60.0 million acre-feet), or roughly four times the mean annual flow at Lee's Ferry, providing a buffer against a significant temporal variability that includes an historic range in annual streamflow at Lee's Ferry of 6 to 28 bcm (5 to 23 maf) (Christensen and Lettenmaier, 2007; USDOI, 2000). 85 % of this storage is in Lakes Powell and Mead, operated by the US Bureau of Reclamation (USBR). Table 1 lists average annual statistics for the 29 sub-basins in which naturalized streamflow is estimated by USBR.

Three principal aquifer systems, as defined by the US Geological Survey (USGS), are included in the basin (Fig. 1; Miller, 1999). The largest of these is the Colorado Plateaus aquifer system, which contains predominantly sandstone whose porosity is low, such that groundwater moves
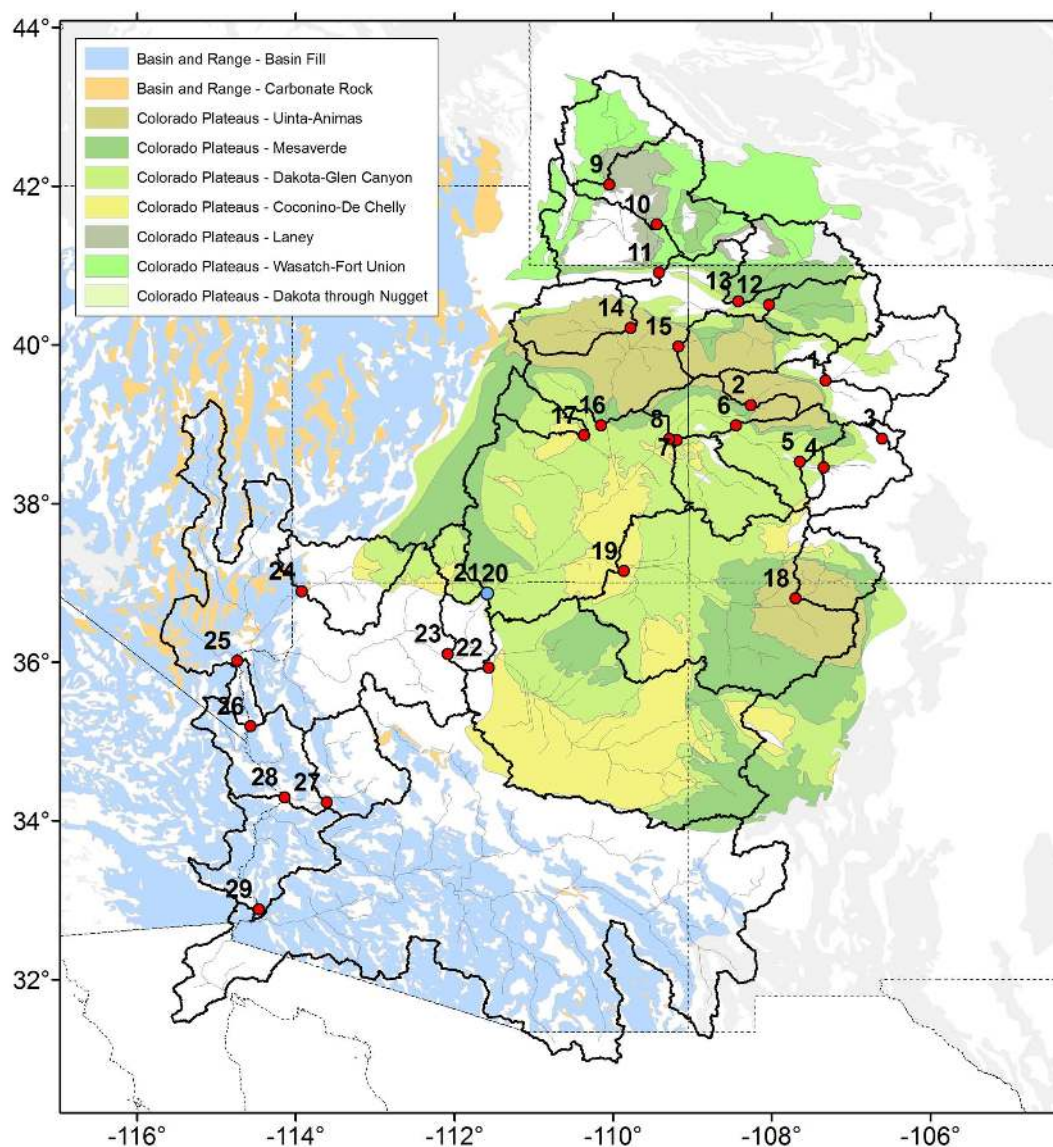
Fig. 1. The Colorado River basin, including the 29 flow locations monitored by USBR, and the uppermost water yielding principal aquifers as given by Miller (1999), Robson and Banta (1995), and Whitehead (1996). Lee's Ferry (station 20) is indicated in blue.

primarily along joints, fractures, and bedding planes. Surficial aquifers in this system include the Uinta-Animas, Mesaverde, Dakota-Glen Canyon, Coconino-De Chelly, Laney, and Wasatch-Fort Union (Robson and Banta, 1995; Whitehead, 1996). The remaining two aquifer systems are located within the Lower Basin and include the Basin and Range basin fill aquifers, generally consisting of unconsolidated gravel, sand, silt, and clay, and the Basin and Range carbonate rock aquifers, consisting of limestone and dolomite (Robson and Banta, 1995). Notably, the Rocky Mountain regions of the basin are not associated with principal aquifer systems.

3 Data and methods

Our experimental approach consisted of two stages. First, we estimated monthly and annual changes in total basin storage and its three main elements – snow water equivalent (SWE), soil moisture, and groundwater – using a combination of physically based hydrologic models (Sect. 3.1), basin-scale water balances (Sect. 3.2), remote sensing data (Sect. 3.3), and baseflow recession analyses (Sect. 3.4), as summarized in Table 2. Second, we analyzed the contribution of groundwater storage to interannual streamflow anomalies using the statistical techniques described in Sect. 3.5.

Table 1. Average annual statistics for the 29 sub-basins over water years 1950–2008. The annual runoff ratio is defined as the ratio of annual runoff to annual precipitation. The mean annual contribution represents the percentage of the total runoff at Imperial Dam (station 29).

		Drainage Area (km^2)	Annual Precip (mm)	Annual Runoff (mm)	Annual Runoff Ratio	Annual Runoff (bcm)	Mean Annual Contrib
1	Colorado R. at Glenwood Springs, CO	11 805	658	211	0.32	2.50	13.0 %
2	Colorado R. near Cameo, CO	20 850	658	202	0.31	4.21	22.0 %
3	Taylor R. below Taylor Park Res., CO	658	726	269	0.37	0.18	0.9 %
4	Gunnison R. below Blue Mesa Dam, CO	8943	615	139	0.23	1.24	6.5 %
5	Gunnison R. at Crystal Res., CO	10 269	618	148	0.24	1.52	7.9 %
6	Gunnison R. near Grand Junction, CO	20 534	566	135	0.24	2.77	14.5 %
7	Dolores R. near Cisco, UT	11 862	475	79	0.17	0.94	4.9 %
8	Colorado R. near Cisco, UT	62 419	558	128	0.23	7.97	41.6 %
9	Green R. below Fontenelle Res., WY	11 085	497	144	0.29	1.60	8.4 %
10	Green R. below Green River, WY	36 260	364	47	0.13	1.71	8.9 %
11	Green R. near Greendale, UT	50 117	367	47	0.13	2.35	12.3 %
12	Yampa R. near Maybell, CO	8832	685	166	0.24	1.46	7.6 %
13	Little Snake R. near Lily, CO	9661	451	57	0.13	0.56	2.9 %
14	Duchesne R. near Randlett, UT	9816	481	95	0.20	0.93	4.9 %
15	White R. near Watson, UT	10 412	472	64	0.14	0.67	3.5 %
16	Green R. near Green River, UT	116 162	411	55	0.13	6.41	33.5 %
17	San Rafael R. near Green River, UT	4217	400	45	0.11	0.19	1.0 %
18	San Juan R. near Archuleta, NM	8443	640	144	0.23	1.22	6.4 %
19	San Juan R. near Bluff, UT	59 570	362	39	0.11	2.34	12.2 %
20	Colorado R. at Lee's Ferry, AZ	289 562	405	61	0.15	17.56	91.6 %
21	Paria R. at Lee's Ferry, AZ	3652	296	6	0.02	0.02	0.1 %
22	Little Colorado R. near Cameron, AZ	68 529	304	3	0.01	0.19	1.0 %
23	Colorado R. near Grand Canyon, AZ	366 744	384	49	0.13	17.98	93.8 %
24	Virgin R. at Littlefield, AZ	13 183	370	16	0.04	0.21	1.1 %
25	Colorado R. below Hoover Dam, AZ/NV	444 703	368	42	0.11	18.58	97.0 %
26	Colorado R. below Davis Dam, AZ/NV	448 847	366	42	0.11	18.88	98.5 %
27	Bill Williams R. below Alamo Dam, AZ	11 999	335	9	0.03	0.10	0.5 %
28	Colorado R. below Parker Dam, AZ/CA	473 193	360	40	0.11	19.07	99.5 %
29	Colorado R. above Imperial Dam, AZ/CA	488 215	353	39	0.11	19.16	100.0 %

Table 2. Summary of storage computation methods.

Method	Sources of Data	Time Period
VIC	Livneh et al. (2013); USBR naturalized streamflow data	1950–2010
VIC-SIMGM	Livneh et al. (2013); USBR naturalized streamflow data	1950–2010
Basin-scale water balance	Livneh et al. (2013); Tang et al. (2009); USBR naturalized streamflow data	2001–2008
Satellite-derived terrestrial water storage change	Swenson and Wahr (2006)	2002–2010
Baseflow recession analysis	Falcone et al. (2010)	Variable

3.1 Hydrologic models

We estimated total basin storage anomalies using two versions of the Variable Infiltration Capacity (VIC) macroscale hydrology model (Liang et al., 1994). VIC is a semi-distributed grid-based model that is typical of LSMs used in numerical weather prediction and climate models (Wood and Lettenmaier, 2006), and has been successfully applied in studies of regions across the conterminous US (CONUS) and worldwide (e.g., Sheffield et al., 2009; Mishra et al., 2010;

Mahanama et al., 2012). Like other LSMs, VIC solves the water and energy balance at each time step, but is distinguished by its parameterization of subgrid variability in soil moisture, topography, and vegetation. In the standard release of VIC (4.0.6 in this study, herein referred to as VIC), no distinction is made between saturated and unsaturated zones in the subsurface.

For our second version (herein referred to as VIC-SIMGM), we modified VIC 4.0.6 to incorporate the SIMple Groundwater Model (SIMGM) of Niu et al. (2007). SIMGM

is one of several recent models that parameterize groundwater as a lumped, unconfined aquifer beneath a multi-layer soil column (e.g., Gedney and Cox, 2003; Yeh and Eltahir, 2005). It is included in Community Land Model (CLM) versions 3.5 (Oleson et al., 2008) and 4.0 (Oleson et al., 2010) and Noah-MP (Niu et al., 2011). In SIMGM, groundwater discharge is parameterized as an exponential function of the water table depth:

$$Q_b = Q_{bmax} e^{-f z_\nabla}, \tag{1}$$

where Q_{bmax} is the maximum groundwater discharge when the water table depth is zero, z_∇ is the water table depth, and f is the decay factor. Groundwater recharge (Q_r) is parameterized by Darcy's law and is positive when water enters the aquifer:

$$Q_r = -K_a \frac{-z_\nabla - (\psi_{bot} - z_{bot})}{(z_\nabla - z_{bot})}, \tag{2}$$

where K_a is the aquifer hydraulic conductivity, z_{bot} is the depth to the bottom of the soil column, and ψ_{bot} is the matric potential of the bottom soil layer. The time rate of change of aquifer storage (dW_a/dt) is then equal to $Q_r - Q_b$, and the water table depth is computed by scaling aquifer storage by the specific yield (S_y). To incorporate SIMGM in VIC, we added a lumped, unconfined aquifer directly to the base of the lowest (third) soil layer and replaced VIC's baseflow scheme with that of SIMGM (Eq. 1). As in CLM, the water table is allowed to move within and between soil layers and the aquifer, in which case Eq. (2) is modified following Niu et al. (2007). Hydraulic conductivity between soil layers is computed as a function of soil texture and water content, whereas hydraulic conductivity of the aquifer decays exponentially with depth from the saturated hydraulic conductivity (K_{sat}) of the lowest soil layer. In VIC-SIMGM, the surface runoff parameterization is identical to that of VIC. Because of differences between VIC and CLM, we do not expect the parameter values in Eq. (1) to match those of Niu et al. (2007).

A limitation of SIMGM is the lack of a direct connection between surface water and groundwater in its parameterization. Because bank storage potentially accounts for a large portion of the interannual hydrologic storage that affects interannual streamflow variations (see, e.g., Meyboom, 1961), this likely has implications for our analysis. A second limitation is the lack of a representation of inter-grid cell (lateral) groundwater flow in SIMGM. Nonetheless, with the exception of some very recent work by Zampieri et al. (2012), most of the land surface groundwater models that have been proposed have similar drawbacks. Thus, we note these limitations but offer results from VIC-SIMGM for comparative purposes.

Meteorological forcing data were gridded from precipitation and maximum/minimum temperature data from National Oceanic and Atmospheric Administration (NOAA) Cooperative Observer stations and wind data from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) Reanalysis Project. These data were derived for the period 1949 to 2010 at a 1/8° resolution using the methods of Livneh et al. (2013), who have extended the data set of Maurer et al. (2002). The first nine months of 1949 were reserved for model spin-up, so that the period of analysis effectively began at the start of water year 1950. Model parameters for VIC were adopted from Christensen et al. (2004, subsequently used by Christensen and Lettenmaier, 2007). For VIC-SIMGM, calibration was performed to monthly naturalized streamflow data for water years 1971–1980 by adjusting Q_{bmax}, f, S_y, d_{mid} (the depth of the middle soil layer), d_{bot} (the depth of the bottom soil layer), and b_{inf} (the infiltration shape parameter in VIC). Naturalized streamflow data for the 29 stations in Table 1 were obtained for the period 1906 to 2008 from USBR (http://www.usbr.gov/lc/region/g4000/NaturalFlow/current.html). For each of the nested sub-basins, calibration was performed in a stepwise fashion, with parameters for the upstream-most sub-basins estimated first. Those parameters were then retained for calibrations to streamflows further downstream.

The VIC-SIMGM implementation required the additional step of first achieving an equilibrium water table depth (WTD) prior to running the simulation. To spin up this term, we ran the model using the observed forcings back-to-back for a period of 2000 yr. The process was expedited by removing the exponential decay function for K_{sat} in the aquifer and simultaneously increasing K_{sat} by up to a factor of 10^6. Results ranged from a depth of 1.3 to 80 m and generally resembled the pattern of annual precipitation for the basin (Fig. 2). The shallow WTDs for some of the more arid locations in the Lower Basin may be related to the model's lack of a parameterization for subgrid topographic variability, which has been shown to enhance interactions between modeled water tables and the land surface, thereby reducing flow paths and increasing subsurface flows (Huang et al., 2008). In any event, our simulations were largely insensitive to this spin-up procedure, probably because those grid cells with the deepest WTDs, which take the longest time to spin up, contribute the least to runoff and experience the smallest changes in hydrologic storage.

Following spin-up, time-varying computations of WTD and groundwater storage were performed for the entire simulation period. Direct validation of simulated WTDs with well observations was complicated (and arguably made infeasible) by the relatively coarse spatial resolution of our model implementation. Because of the great spatial variability in WTDs, we, like Niu et al. (2007), did not expect our simulated WTDs to be representative of what amount to (nearly) point observations. We nonetheless argue that the model is appropriate for evaluation of changes in groundwater storage, and hence used the analyses that follow for this purpose.
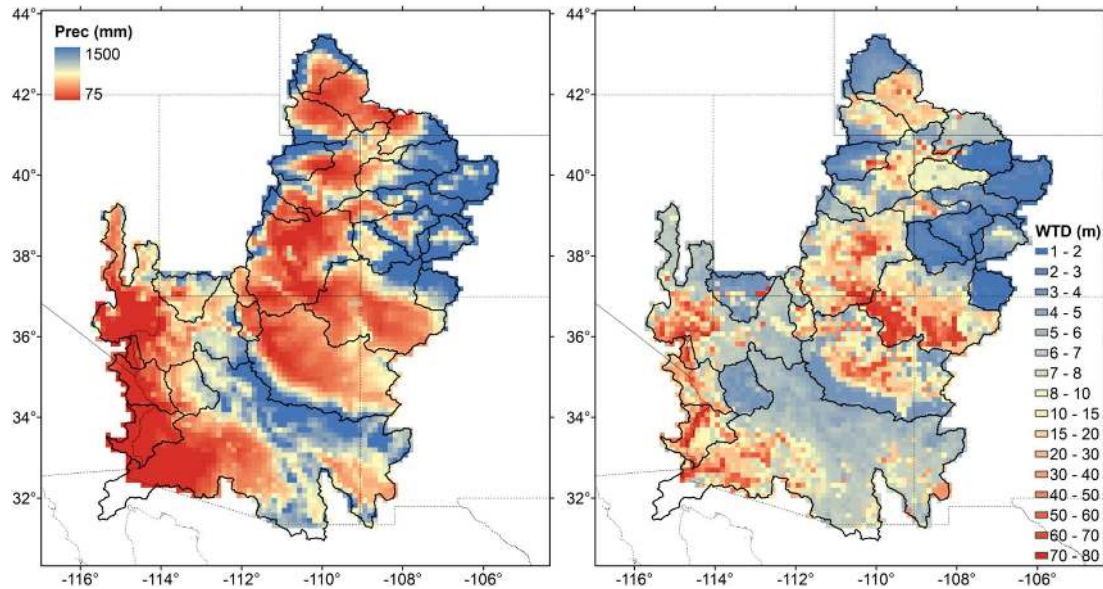
Fig. 2. Average annual precipitation over water years 1950–2008 (left) and equilibrium water table depth as simulated by SIMGM (right).

3.2 Basin-scale water balance

A central question in our study is how well VIC captures the variability in hydrologic storage for the Colorado River basin. This is not unlike a more traditional problem in water resources, which is how well a reservoir's stage-storage relationship captures the variability in its contents. Langbein (1960) addressed such a question for Lake Mead by carefully comparing changes in reservoir volumes as determined by a stage-storage relationship from a recent bathymetric survey with changes in volumes as calculated through a water budget of observed inflows and outflows. He found that gains in volume as derived from the water budget exceeded those derived from the stage-storage relationship in wet years, and losses in volume as derived from the water budget exceeded those derived from the stage-storage relationship in dry years, with the magnitudes of the residuals proportional to the changes in volume. He attributed these residual quantities to a "hidden" storage term that was neglected in the stage-storage relationship, namely, bank storage in the sediment of the reservoir, which he estimated at roughly 3 million acre-feet, as compared to the lake's usable capacity of 27 million acre-feet at the time. In a similar study, Murdock and Calder (1969) estimated about 6 million acre-feet of bank storage for Lake Powell, which also has a usable capacity of 27 million acre-feet.

In this study, we adopted a similar approach to assess any "hidden" components in the modeled storage quantities by comparing them with changes in storage that are derived exclusively from observational data. Rather than a single reservoir, however, we adopted each of the 29 sub-basins as our control volumes, and computed changes in storage using the

hydrologic continuity equation:

$$\Delta S = P - Q - \mathrm{ET}. \tag{3}$$

We used gridded precipitation data (at a $1/8°$ spatial resolution) and naturalized streamflow taken from the monthly USBR data described in Sect. 3.1. For ET, we used a satellite-based product from MODIS (Tang et al., 2009), aggregated from $0.05°$ to a $1/8°$ resolution. The MODIS-based product is one of several recent ET data sets derived primarily or entirely from remote sensing data (e.g., Ferguson et al., 2010; Zhang et al., 2010). These data, arguably, are the nearest alternative to ET observations available at this scale.

For the pilot study area of the Klamath River basin, Tang et al. (2009) found daily ET biases of less than 15 % when compared with ground flux tower observations and Landsat-based estimates, with a tendency for the MODIS-based product to underestimate seasonal ET. They noted that the algorithm was most effective over areas containing a substantial diversity in vegetation types. In an analysis of a similar product, Ferguson et al. (2010) found that satellite-based ET was also biased low when compared with VIC-simulated ET for the continental US, except for the Colorado and Great Basins, where it was biased high. They suggested that the most likely explanation for the high bias was the lack of a constraint of soil water availability for the remote sensing product.

3.3 Satellite-derived terrestrial water storage change

As an additional basis for comparison, we analyzed estimates of terrestrial water storage change (TWSC) from the GRACE satellite mission since its 2002 launch. Despite a relatively coarse effective spatial resolution of several hundred

kilometers, GRACE data have demonstrated utility for quantifying changes in hydrologic storage in a number of recent studies. Syed et al. (2008) found good agreement between GRACE-derived TWSC and Global Land Data Assimilation System simulations at global and continental scales. Strassberg et al. (2009) found that GRACE-derived data were highly correlated with in situ soil moisture and groundwater observations for the High Plains aquifer, and Grippa et al. (2011) showed that GRACE data adequately reproduced the interannual variability of water storage estimated by nine LSMs over West Africa. In a study of nine major US river basins, Gao et al. (2010) found that GRACE data tended to underestimate TWSC when compared with VIC simulations, although they noted that errors tended to be of smaller magnitude than those for satellite-based ET and precipitation.

The GRACE data sets used here were processed by the Center for Space Research (CSR) at the University of Texas and were filtered to remove spatially correlated errors that result in north–south data "stripes" (Swenson and Wahr, 2006). Data were provided at a 1° spatial resolution for the period 2002–2010 and represent "equivalent water thickness" as computed from observations collected continuously over monthly intervals.

3.4 Baseflow recession analysis

We inferred storage changes for subsets of the domain using baseflow recession analysis. For this we used daily streamflow data from 72 "reference-quality" gages in the GAGES-II database (http://water.usgs.gov/GIS/metadata/usgswrd/XML/gagesII_Sept2011.xml; Fig. 3). This is an update to a compilation of all USGS stream gages in CONUS that are either currently active or have at least 20 yr of complete-year flow records since 1950, and for which watershed boundaries can be reliably delineated (Falcone et al., 2010).

We adopted two separate approaches for our recession analysis. The first, more conventional, approach utilizes the classical form of the recession equation as proposed by Barnes (1939) and Maillet (1905):

$$Q = Q_0 k_r^t, \tag{4}$$

where Q_0 is the streamflow at some arbitrary time $t = 0$ and k_r is the recession constant. For each of the 72 GAGES watersheds, we derived a single recession constant k_r using the semi-automated methodology of RECESS, available for download from http://water.usgs.gov/software/lists/groundwater and described by Rutledge (1998).

Our second approach is based on the work of Brutsaert and Nieber (1977), who described a family of recession curves by the derivative of the nonlinear equivalent of Eq. (4):

$$dQ/dt = -aQ^b, \tag{5}$$

where a and b are constants. Vogel and Kroll (1992) further showed that b can be assumed to be one, in which case only
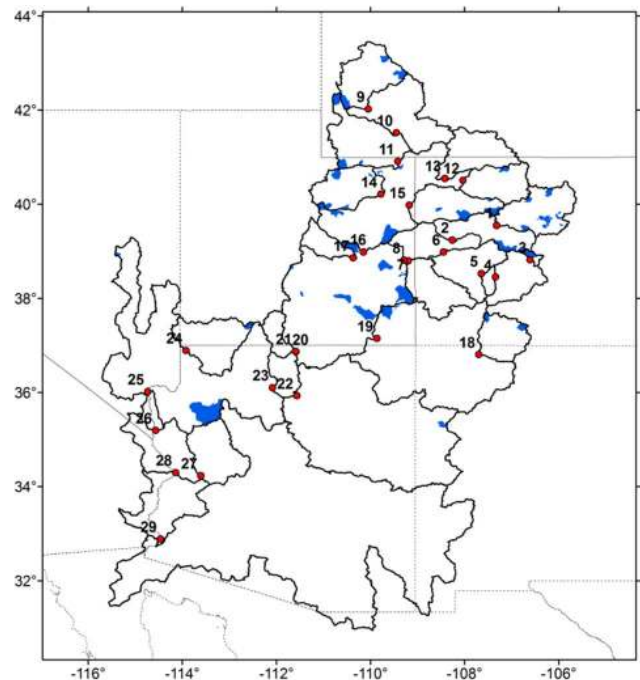

Fig. 3. The locations of the 72 "reference-quality" watersheds (in blue) used in the baseflow recession analysis. The 29 USBR gauges and sub-basins are shown for reference.

a needs to be estimated from streamflow data. The reciprocal of a has been labeled as the recession timescale τ (Eng and Milly, 2007), which is thus equal to $-Q/(dQ/dt)$ and can be used to relate streamflow Q to basin water storage S. For each streamflow record, we identified recession periods using the same criteria as in RECESS, which assumes discharge originates entirely from groundwater at $N = A^{0.2}$ days following each streamflow peak, with A representing the drainage area in square miles (Linsley et al., 1982). We then followed the methodology of Kirchner (2009) and Krakauer and Temimi (2011) to compute τ from streamflow observations during these recession periods and develop S-Q curves for each watershed. Values of S were derived for the beginning of each water year (1 October) as those corresponding to the minimum streamflow during the 30-day window of 16 September to 15 October, thus assuring at least a high probability that this streamflow was baseflow-dominant, even if it did not fall during a recession period in the strict sense. We then computed annual storage changes for comparison with our other estimates.

Kirchner (2009) further argued that storage–discharge relationships derived from periods of recession could be assumed for other parts of the hydrograph, as evidenced by the ability to invert these relationships to successfully infer (P-ET) from streamflow observations for his two test basins in the United Kingdom. Krakauer and Temimi (2011) later used this assumption to estimate monthly storage as the mean of hourly S derived from observed Q, finding that it did not
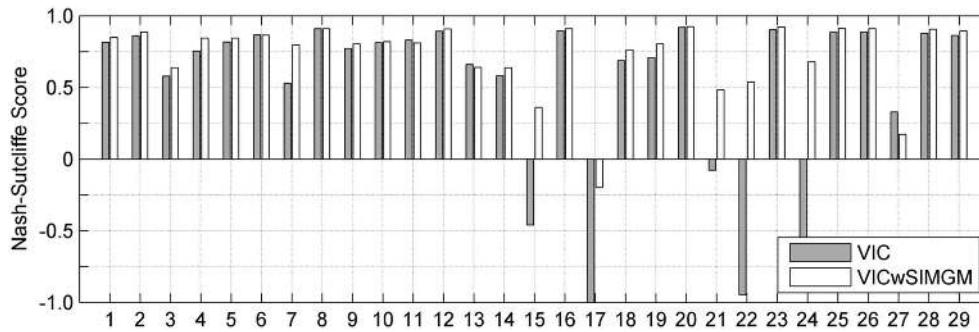
Fig. 4. Nash–Sutcliffe efficiency scores for VIC and VIC-SIMGM in each of the 29 sub-basins.

necessarily hold in 61 small watersheds across the US. Bearing these results in mind, we tested this approach to also infer monthly storage changes from our daily streamflow data.

3.5 Statistical analysis

We performed several first-order statistical analyses on the estimates of storage and storage change. Recognizing that a key issue in our study's main objective relates to the contribution of carryover storage from the previous water year, we proposed the naïve hypothesis that interannual hydrologic storage contributes most to streamflow during years of drought, much like a reservoir is drained to compensate for dry conditions. We tested this hypothesis by evaluating the relationship between change in total water year storage and water year streamflow volume, assuming that, except for losses to evapotranspiration, any negative change in water year storage can be considered a contribution to streamflow from the previous water year. We also examined relationships between water year storage change and 1 October storage anomaly, and water year storage change and previous water year streamflow volume, as a basis for comparison.

We then explored the utility of groundwater estimates for seasonal streamflow forecasts, which are typically issued for the target period April–July in the Colorado River basin by the NWS Colorado Basin River Forecast Center. Here we adapted the dimensionless parameter κ, which was introduced by Mahanama et al. (2012) as the ratio of the standard deviation of total basin storage (at a forecast lead of zero) to that of precipitation during the forecast target period. As such it is essentially a comparison between the known and unknown water volumes that determine streamflow, providing an approximate measure of the predictability that can be derived solely from IHCs. Using ensemble streamflow forecasts for periods ranging from one to six months, Mahanama et al. (2012) and Shukla and Lettenmaier (2011) found first order relationships between κ and forecast skill for 23 basins and 48 hydrologic sub-regions across CONUS, respectively. We extended the κ concept to measure the predictive capacity of IHCs at leads greater than zero, and also to compare the variability of individual storage terms in addition to total

basin storage:

$$\kappa = \sigma_w / \sigma_p, \tag{6}$$

where σ_w is the standard deviation of groundwater, soil moisture, SWE, or various combinations thereof on the forecast issue date, and σ_p is the standard deviation of precipitation up to and including the forecast target period. We then developed simple forecasts of April–July streamflow volume via multiple linear regression with the basin-averaged storage terms, and examined the co-variability of the skill (R^2) of these forecasts and our estimates of κ.

4 Results

We present our findings in a similar fashion to the methods described above. Sections 4.1 to 4.3 provide results from the model simulations, satellite data comparisons, and recession analyses in order to establish a most plausible set of storage change estimates. Section 4.4 analyzes these estimates to assess the contribution of groundwater storage to interannual streamflow anomalies.

4.1 Model performance

Figure 4 shows Nash–Sutcliffe scores for both VIC and VIC-SIMGM in each of the 29 sub-basins. The performance of the two models was quite similar. At Lee's Ferry, scatter plots of simulated vs. observed annual streamflows appeared almost identical, showing good agreement between the two terms over the study period. Results for other sub-basins were comparable. For VIC-SIMGM, basin-averaged WTDs exhibited a clear seasonality and ranged from an average of 1.2 m (sub-basin 18) to 27.5 m (sub-basin 19).

Details on subsurface storage simulations are given for Lee's Ferry and its largest headwater basin (Glenwood Springs) in Fig. 5. As noted by Niu et al. (2007) for CLM, VIC-SIMGM resulted in bottom soil layers that were wetter by volumetric water content than in VIC. Differences in wetness were proportional to WTDs and were most evident during the snow accumulation season. On the other hand,
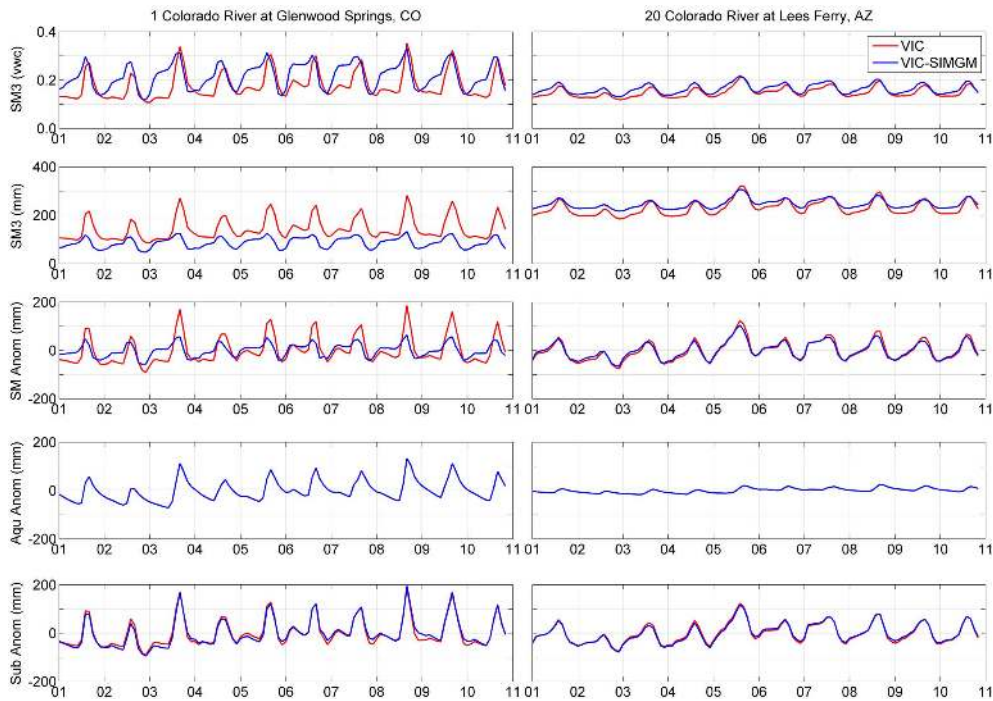
Fig. 5. Time series of the subsurface storage terms (from top to bottom): third layer soil moisture by volumetric water content, third layer soil moisture in mm, total soil moisture anomaly, aquifer storage anomaly, and total subsurface storage anomaly.
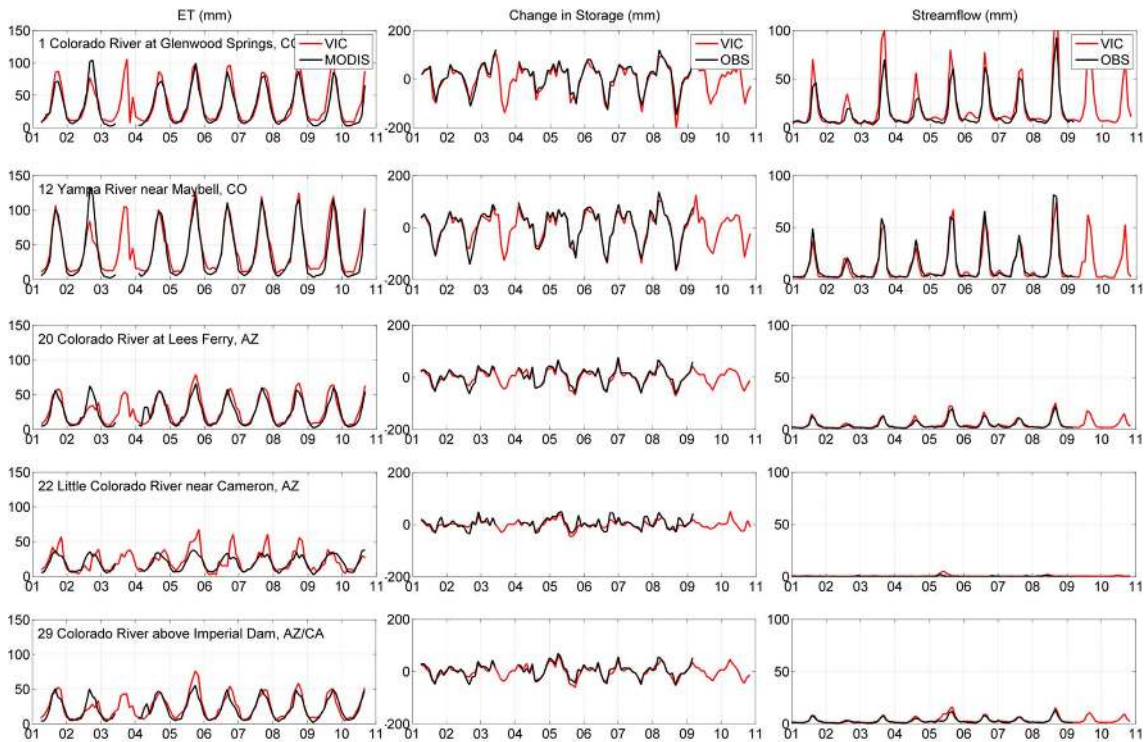

Fig. 6. Comparisons between MODIS-derived and VIC ET (left), results of the basin-scale water balance (middle), and comparisons between simulated and observed streamflow (right) for selected sub-basins.
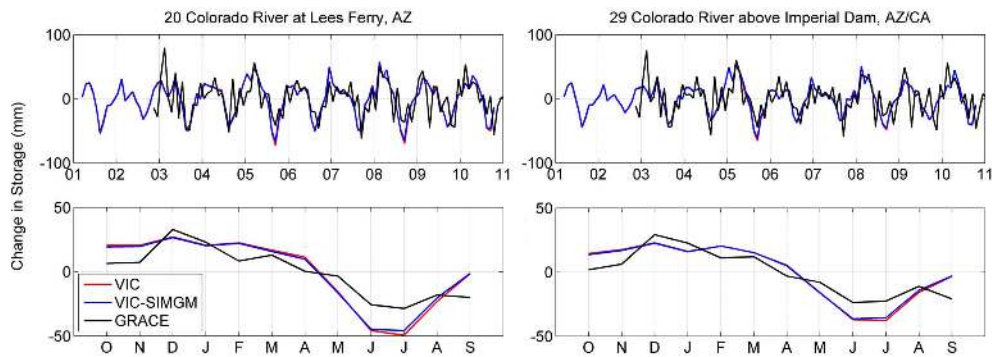
Fig. 7. Comparisons between VIC-simulated and GRACE-derived changes in storage.

bottom layer soil moisture volumes were generally smaller for VIC-SIMGM in sub-basins with shallow WTDs, and total soil moisture anomalies were likewise smaller for VIC-SIMGM. These differences, however, were approximately equal to aquifer storage anomalies, so that total subsurface storage anomalies were roughly the same for both models. The mechanism behind these differences is apparent upon examination of the soil layer depths – for those headwater sub-basins with the shallowest WTDs, calibration resulted in soil columns that were shallower in VIC-SIMGM than in VIC in order to conserve mass. Differences between the models in total basin storage were likewise negligible.

Comparisons of flux terms for the two models were generally unremarkable, despite different baseflow parameterizations. While the wetter soil moisture profiles in VIC-SIMGM would be expected to increase runoff efficiencies, calibration resulted in runoff efficiencies that were comparable to those of VIC. In VIC-SIMGM, baseflow constituted a slightly higher percentage of runoff for regions of shallow WTDs, and the annual recession of the simulated hydrograph was noted to typically better match that of the observed hydrograph. ET was also somewhat lower in VIC-SIMGM than in VIC, a result that is consistent with other studies of LSM groundwater parameterizations (e.g., Liang et al., 2003).

Modeled storage terms were verified against observations to the extent possible. Measured SWE data were obtained for all active NRCS snow courses and SNOTEL stations within the basin from http://www.wcc.nrcs.usda.gov/reportGenerator (116 snow courses and 186 SNOTEL sites total). In each sub-basin, 1 April SWE was computed by averaging the available observations for each water year. VIC-derived estimates were similarly obtained by averaging SWE simulated at the grid cells in which the observations were located. Direct comparisons of these averages generally showed significant differences in magnitude, with observed estimates consistently higher than simulated estimates, a result that is not surprising given the mismatch in scale from point to grid cell. Standardized Z-scores of these estimates, however, compared more favorably.

4.2   Satellite-derived storage changes

As a precursor to performing the basin-scale water balance, we compared MODIS-derived estimates of ET with VIC estimates of ET for each sub-basin. As shown in Fig. 6 (left), the two estimates generally matched well over the period of MODIS observations. Some slight discrepancies can be seen, such as a tendency for VIC to estimate more ET than MODIS for winter months in sub-basin 12 and for summer months in sub-basin 22. The most consistent difference between the two products occurs in water year 2002, for which MODIS estimates substantially more ET than VIC. Since 2002 was an extremely dry year (driest or second driest for most sub-basins), this is likely a result of the lack of a constraint of soil water availability for the MODIS product (see Sect. 3.2).

Monthly changes in storage as derived from the basin-scale water balance and VIC are shown in the center column of Fig. 6, and for comparative purposes, simulated and observed streamflow hydrographs are shown for the same time period on the right. The two storage estimates match quite well for all sub-basins, except for water year 2002 due to the previously noted discrepancy in ET. Following the work of Langbein (1960), annual changes in VIC-simulated storage were plotted against the residuals of the two quantities (not shown) to infer any element of hydrologic storage that is not captured by VIC (see Sect. 3.2). Despite the good agreement at a monthly time step, however, no consistent relationship was found at an interannual time scale, due in part to the inadequacy of the sample size.

Figure 7 compares simulated and GRACE-derived changes in storage for the sub-basins of Lee's Ferry (20) and Imperial Dam (29). Reasonable agreement between the estimates can be seen, although those derived from GRACE are noticeably noisier than those simulated by VIC. Interestingly, the GRACE data generally indicate more modest changes in storage than VIC, particularly during the summer months of June and July; VIC-SIMGM only slightly reduces this difference. Note that GRACE data represent all changes in water storage, including surface water in lakes and reservoirs. Based on observations of interannual storage change
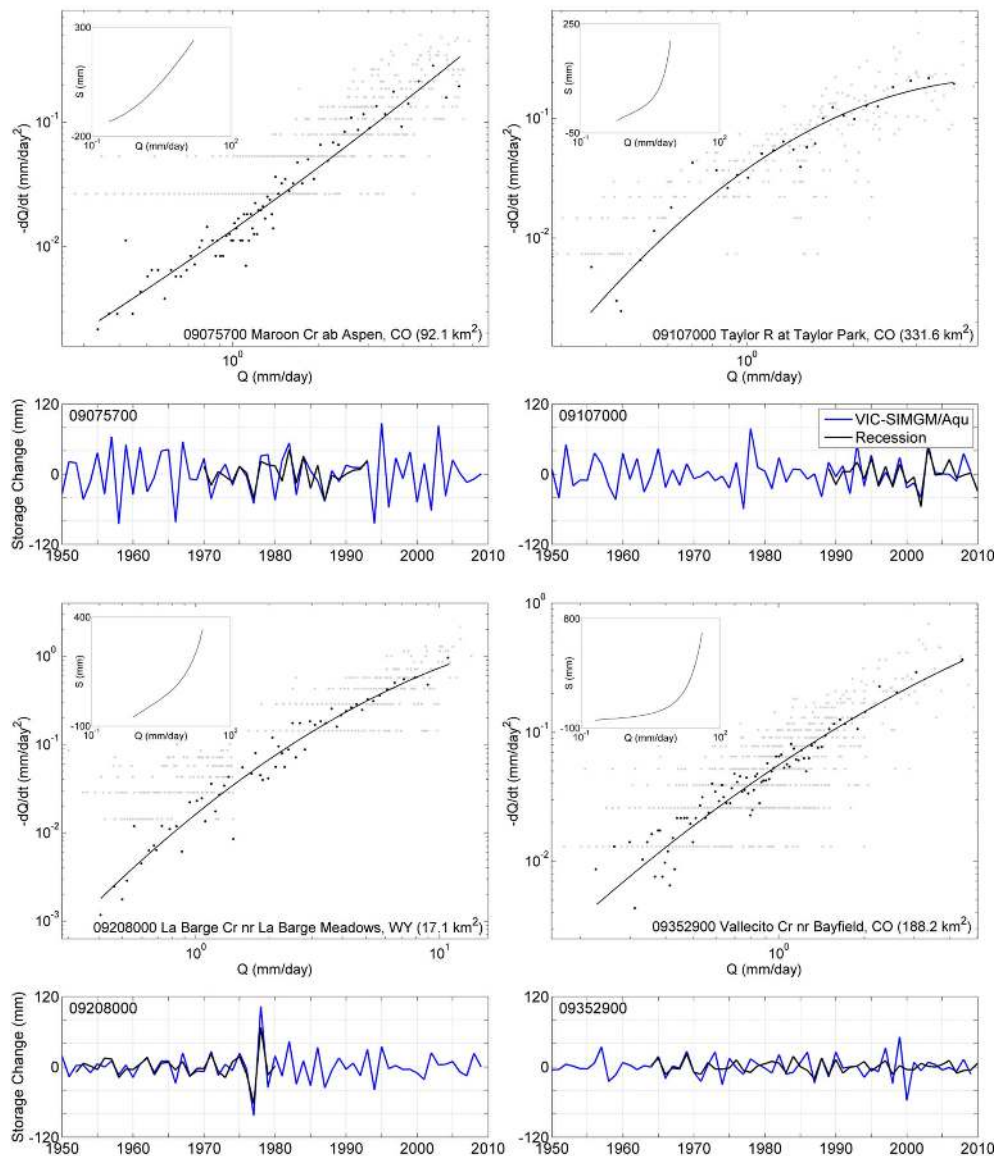
Fig. 8. Recession plots and associated storage functions (insets) for four reference watersheds. Annual changes in groundwater storage as derived from storage functions are compared with those derived from VIC-SIMGM below their respective recession plots. Watersheds are located in sub-basin 2 (09075700), sub-basin 3 (09107000), sub-basin 9 (09208000), and sub-basin 18 (09352900), with drainage areas provided in parentheses.

in Lakes Mead and Powell, we estimated the effect of these changes to be a maximum of 15 mm averaged basin-wide. Given that most of this surface water storage change occurs in the Lower Basin, while most change in hydrologic storage occurs in the Upper Basin, this effect appears to be minimal in our area of interest.

4.3 Recession analysis

Baseflow recession constants as derived from RECESS were generally consistent across watersheds. Most results for watersheds in headwater basins fell in the upper part of the typical range at about 0.96–0.99, with only one watershed in sub-basin 9 indicating a recession constant smaller than 0.95. Recession constants for watersheds in the lower part of the Upper Basin were somewhat lower, with a few near 0.85.

Recession plots of $-\mathrm{d}Q/\mathrm{d}t$ vs. Q are given for four representative watersheds in Fig. 8. The gray dots in these plots represent all recession observations for each watershed, while the black dots are binned averages following the methodology of Kirchner (2009). Recession timescales (τ) derived from the fitted curves tended to increase with decreasing streamflow but generally were in the 45 ± 15 day range cited by Brutsaert (2008); for the watersheds shown
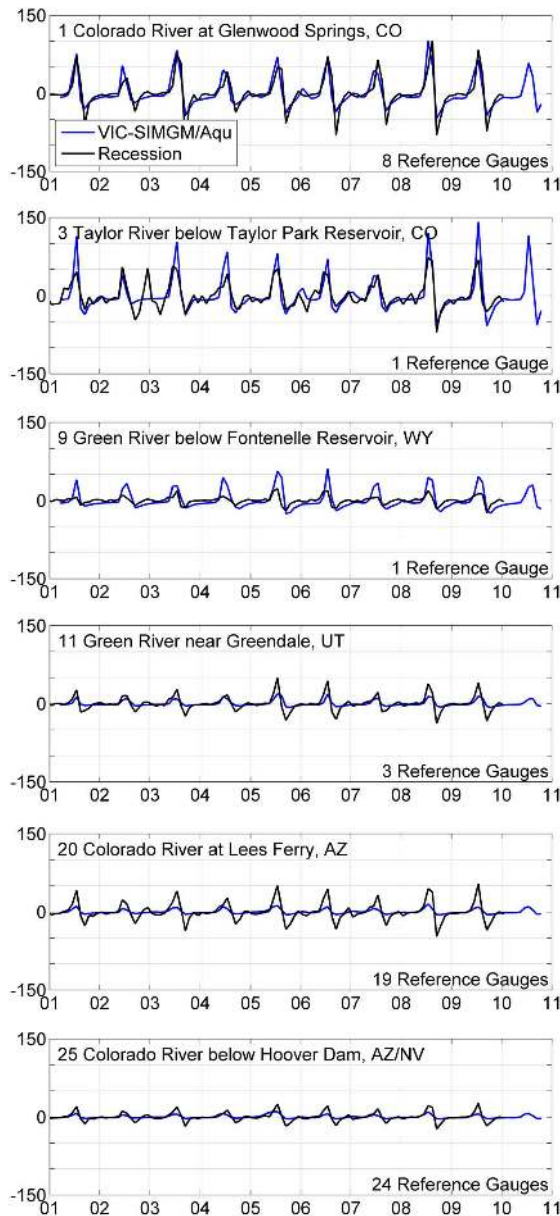
Fig. 9. Comparison of monthly change in groundwater storage (y-axis, in mm) as derived from VIC-SIMGM and recession analysis. The number of reference watersheds used for each sub-basin is provided at lower right.

Fig. 10. Scatter plots for Lee's Ferry of **(a)** annual storage change vs. annual streamflow, **(b)** annual storage change vs. 1 October storage anomaly, and **(c)** annual storage change vs. previous water year streamflow. Annual storage changes were calculated between 1 October and 1 October of the following year. Red, yellow, and blue circles denote dry, normal, and wet water years, respectively.

in Fig. 8, for example, τ ranged from ~ 20 days (gauge #09352900) to ~ 80 days (gauge #09075700) at a Q of $1\,\mathrm{mm\,day^{-1}}$. Storage functions derived from these recession plots are shown in the inset graphs of Fig. 8 (note that the y-axis on these graphs is labeled with respect to an arbitrary datum), and annual changes in groundwater storage estimated from these functions are shown below their respective recession plots. For comparative purposes, we also show changes in VIC-SIMGM aquifer storage for the grid cell nearest the centroid of each watershed, which typically covered an area
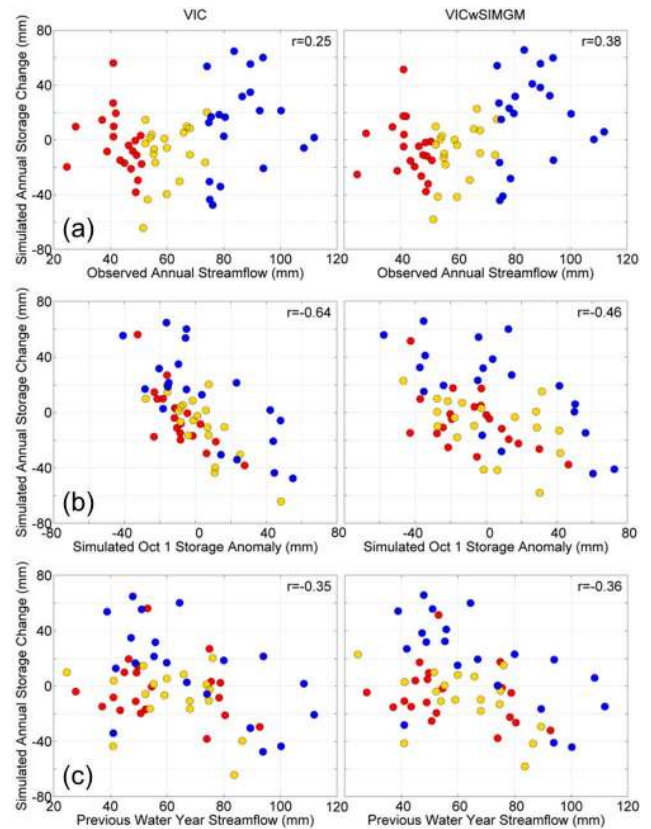
about the same size (and no larger than 2 or 3 grid cells at most). For the watersheds shown here, the ranges of variability in the two estimates are roughly similar, with storage changes for gauge #09107000 and 09208000 matching particularly well. For some of the other watersheds, we noticed a tendency for the simulated groundwater storage to have greater interannual variability than the storage inferred from the recession analysis.

Figure 9 presents results from the monthly storage analysis, which assumed that storage–discharge relationships derived from periods of recession could be applied to other parts of the hydrograph as well. To obtain sub-basin-wide estimates of storage change, storage time series were developed for all watersheds with complete records for 2001–2010, and storage averages (weighted by watershed area) were then computed where possible (about half of the sub-basins). Somewhat remarkably, these estimates matched up well with those derived from VIC-SIMGM for sub-basins 1, 3, and (to a lesser extent) 9. The agreement in 3 is perhaps less surprising given that the reference watershed (gauge
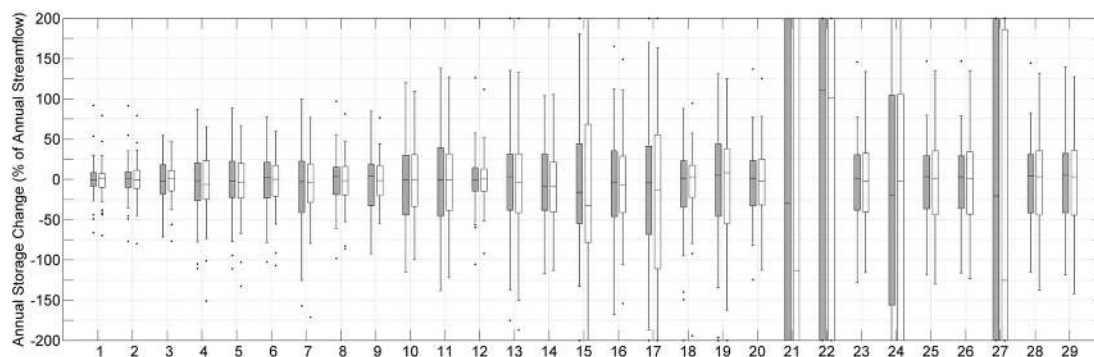
Fig. 11. Ranges of simulated annual storage change as a percentage of annual streamflow observations for VIC (gray boxes) and VIC-SIMGM (white boxes). The central mark in each box indicates the median, and the edges of the box are the 25th and 75th percentiles. Whiskers extend to a maximum of ~ 2.7 standard deviations from the median, and outliers are indicated by dots.

#09107000, which is also shown in Fig. 8) accounts for roughly half the area of the sub-basin, but the same cannot be said for sub-basins 1 and 9. For sub-basins further downstream (11, 20, and 25), recession-derived storage changes tend to vary more greatly than those derived from VIC-SIMGM, which can probably be attributed to the headwater bias of the reference watersheds.

4.4 Statistical analysis

As tests of the hypothesis that interannual hydrologic storage contributes most to streamflow during dry years, comparisons among simulated storage estimates and annual streamflow observations are shown for Lee's Ferry in Fig. 10; for these plots streamflow records were divided into terciles of dry (red circles), normal (yellow circles), and wet (blue circles) water years. As shown in the top two plots (Fig. 10a), correlations between simulated total water year storage change and observed water year streamflow volume are weak at best, with a slight tendency to lose storage in dry years that is modestly more pronounced in VIC-SIMGM. Correlations between simulated annual storage change and simulated 1 October storage anomalies (Fig. 10b), on the other hand, are somewhat stronger, which is fairly intuitive – what goes up must come down and vice versa. We therefore hypothesized that simulated annual storage changes were related to observed streamflow volumes from the previous water year (Fig. 10c), but again found only weak correlations, due in part to the fact that previous water year streamflow volumes are poor indicators of the previous water year's storage change (and hence 1 October storage anomalies) to begin with, as shown in Fig. 10a. Results for other sub-basins were comparable, as were comparisons between simulated storage estimates and simulated streamflow volumes. Thus, annual streamflow volumes appear to bear little relation to interannual hydrologic storage, and whether interannual hydrologic storage (and consequently, interannual groundwater storage) contributes to streamflow appears to be more a function of the

initial storage conditions, being more or less equally likely in a wet or dry year.

Figure 11 provides an alternative perspective on this issue by comparing the range of simulated annual storage change as a percentage of that year's streamflow observation for each of the 29 sub-basins. Expressed in this way, the ranges are smallest for sub-basin 1, the single largest contributor of runoff among the headwater basins, and are minimal for similarly significant sub-basins 2, 8, and 12. On the other hand, ranges are largest for Lower Basin sub-basins that yield much smaller runoff volumes. Thus, interannual hydrologic storage appears of least importance where streamflow matters most, and most where streamflow matters least. As noted in Sect. 4.1, it is difficult to discern any consistent difference between storage changes for VIC and VIC-SIMGM; the range of variability is wider for VIC in sub-basins 7 and 11 and for VIC-SIMGM in sub-basins 15 and 17, but by and large the two are comparable. At Lee's Ferry, both suggest a lower quartile of about $-25\,\%$ and an upper quartile of about $+25\,\%$.

Results from the κ analysis are shown in Fig. 12. The top two plots compare skill (here expressed as R^2) for regression forecasts of an April–July target period against κ at lead zero, with each dot representing a different sub-basin, and each color a different simulated storage term or combination of terms. The largest dots in each plot are for Lee's Ferry. For VIC, SWE and soil moisture exhibit comparable estimates of κ, although κ is more variable for SWE across sub-basins, and forecasts based on SWE are noticeably more skillful. As expected, total storage results in values of R^2 and κ that are greater than for either of the individual storage terms. These patterns are largely replicated for VIC-SIMGM, although soil moisture exhibits marginally lower values of κ than for VIC (as discussed in Sect. 4.1), despite seemingly no loss in forecast skill. Aquifer storage has the lowest κ of all, with (somewhat disproportionately) little to no forecast skill. A first order relationship between R^2 and κ is not particularly
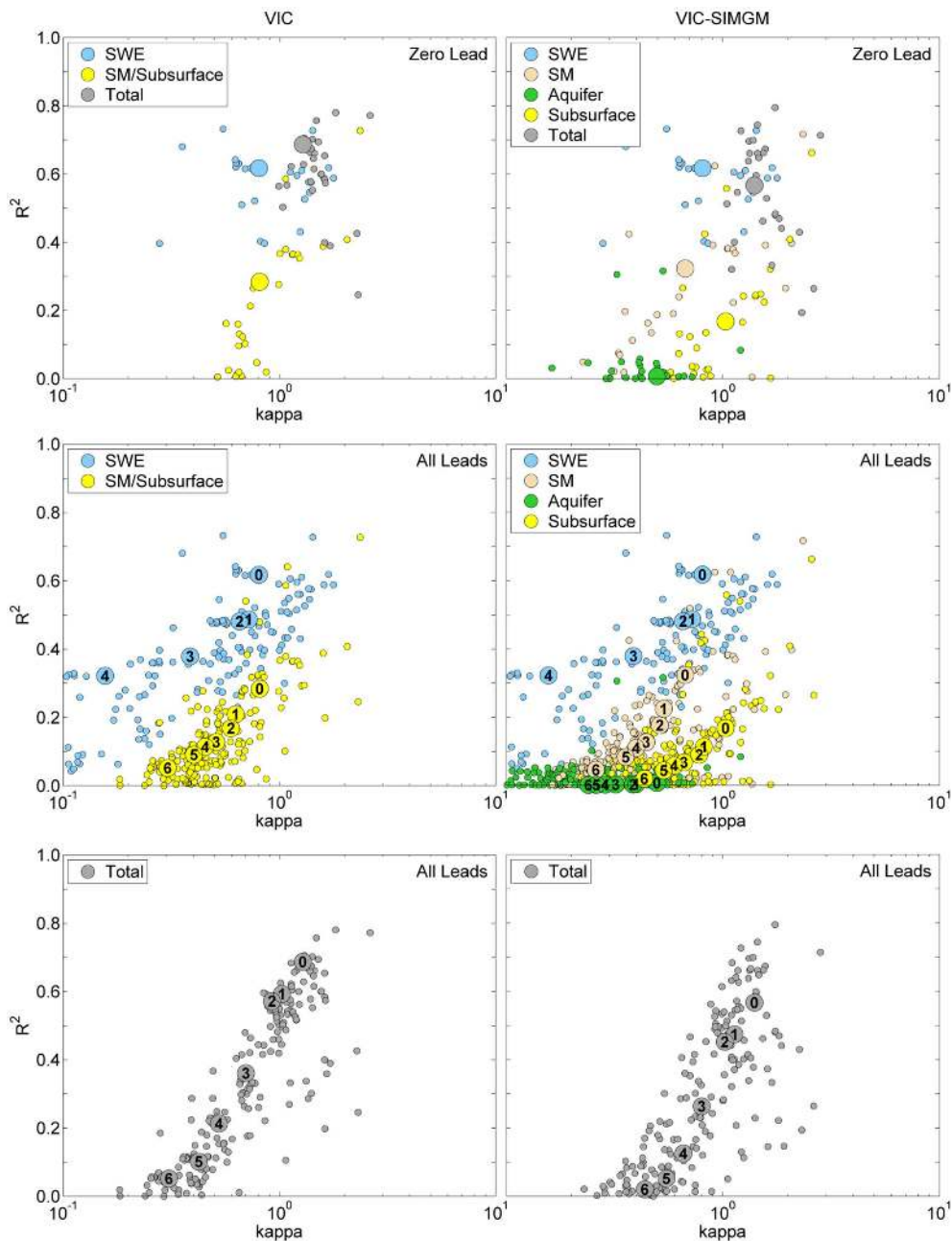
Fig. 12. Scatter plots of forecast skill vs. kappa for various storage term combinations, reflecting an April–July target period. The larger dots in each plot are for Lee's Ferry, with the numbers inside indicating the lead time in months.

obvious from these plots, perhaps due to the small sample size.

The middle two plots in Fig. 12 compare individual simulated storage terms at leads up to six months (i.e., forecasts issued on 1 October), with the numbers inside the larger dots denoting the lead time for Lee's Ferry. Similar patterns can again be seen: SWE exhibits the highest skill, VIC-SIMGM soil moisture exhibits lower values of κ than for VIC, and aquifer storage exhibits the lowest values of κ with negligible

forecast skill. Though somewhat noisy, a first order relationship between R^2 and κ is here more apparent, with the odd exception of aquifer storage. For the bottom two plots, which show results for total basin storage, this first order relationship is clearly visible, demonstrating that the κ concept can be extended to lead times beyond zero, at least where total basin storage is concerned.

The apparently counterintuitive result for aquifer storage merited additional investigation. Correlations between

basin-averaged aquifer storage and April–July streamflow were found to be lower than for other storage terms, but the reason was not obvious. Autocorrelation functions were strong and, somewhat logically, higher than for other terms. We tested whether different combinations of grid cells resulted in greater predictive capacity than the collective basin average using a search routine based on principal components regression (Rosenberg et al., 2011). Skill improved only marginally, however, and negligibly in relation to the improvement found by adopting this approach for other storage terms.

5   Summary and conclusions

Our analysis quantitatively assessed the significance of groundwater storage to interannual streamflow anomalies in the Colorado River basin. We compared several estimates of total basin storage change and found that VIC simulations yield similar results regardless of whether a groundwater representation is included. We also found that these estimates compare favorably with those obtained from observation-based basin-scale water balances and GRACE measurements. Further, we found that changes in VIC-SIMGM-simulated groundwater storage were similar to those derived from a baseflow recession analysis.

Assessments of the co-variation between simulated annual storage change and water year streamflow volumes revealed essentially no relationship between these two terms for any of the 29 sub-basins. Similarly, interannual hydrologic storage accounted for only a small percentage of annual streamflow in the headwater sub-basins from which most of the basin's runoff originates. Simulated estimates of groundwater storage exhibited less variability and weaker seasonal streamflow predictive skill than either SWE or soil moisture at every lead time. Thus, we conclude that groundwater storage does not provide a significant contribution to interannual streamflow anomalies in the Colorado River basin.

The limited role of groundwater storage in Colorado River streamflow anomalies is most likely due to the subsurface geology of the basin. The sandstone aquifer system underlying the majority of the Upper Basin is largely disconnected from the streamflow generation process, particularly in the central, semiarid rangeland where water tables are deep. For headwater sub-basins where greater groundwater–streamflow interaction might be expected, our results (Sect. 4.4) suggest that storage conditions at the beginning of the water year serve to constrain the range of subsequent changes in storage. Thus, while low annual precipitation should partition into negative anomalies for both surface runoff and groundwater recharge, the smaller value of recharge simply results in a smaller value of groundwater discharge if the water table is already low. It is worth mentioning that, since part of our analysis was based on the extremely dry decade of the 2000s, our findings might differ somewhat for a decade of greater interannual variability, in which case initial conditions may be less influential and groundwater contribution to streamflow might be somewhat larger.

The implications of these results are noteworthy for both operational and long-term planning purposes. Operationally, they suggest that current statistical and ensemble-based water supply forecasts, neither of which account for groundwater conditions, are likely not detrimentally affected by this omission in the Colorado River basin. With respect to a longer timeframe, the results imply that there is little dependence of one year's discharge on that of the previous year, an issue that has been cited as potentially helping to reconcile modeled projections of climate-change-induced reductions in Colorado River streamflow (Hoerling et al., 2009). Nonetheless, the contribution of groundwater storage to interannual streamflow anomalies on a global scale is an important and relatively unexplored issue. The methods presented herein can be used to evaluate this issue for other locales.

References

Barnes, B. S.: The structure of discharge-recession curves, Trans. Am. Geophys. Union, 20, 721–725, 1939.

Barnett, T. P. and Pierce, D. W.: Sustainable water deliveries from the Colorado River in a changing climate, Proc. Natl. Acad. Sci., 106, 7334–7338, 2009.

Boardman, H. P.: The effect of soil-absorption on snow-survey forecasting of stream-flow, 4th Western Snow Conference, Pasadena, California, 31 January 1936, 534–537, 1936.

Brutsaert, W.: Long-term groundwater storage trends estimated from streamflow records: Climatic perspective, Water Resour. Res., 44, W02409, doi:10.1029/2007WR006518, 2008.

Brutsaert, W. and Nieber, J. L.: Regionalized drought flow hydrographs from a mature glaciated plateau, Water Resour. Res., 13, 637–643, 1977.

Christensen, N. S. and Lettenmaier, D. P.: A multimodel ensemble approach to assessment of climate change impacts on the hydrology and water resources of the Colorado River Basin, Hydrol. Earth Syst. Sci., 11, 1417–1434, doi:10.5194/hess-11-1417-2007, 2007.

Christensen, N. S., Wood, A. W., Voisin, N., Lettenmaier, D. P., and Palmer, R. N.: Effects of climate change on the hydrology and water resources of the Colorado River Basin, Clim. Change, 62, 337–363, 2004.

Church, J. E.: Principles of snow surveying as applied to forecasting stream flow, J. Agric. Res., 51, 97–130, 1935.

Clyde, G. D.: Soil-moisture studies as an aid in forecasting runoff from snow-cover, 8th Western Snow Conference, Portland, Oregon, 19 April 1940, 871–873, 1940.

Day, G. N.: Extended streamflow forecasting using NWSRFS, J. Water Resour. Plan. Manage., 111, 157–170, 1985.

Eng, K. and Milly, P. C. D.: Relating low-flow characteristics to the base flow recession time constant at partial record stream gauges, Water Resour. Res., 43, W01201, doi:10.1029/2006WR005293, 2007.

Falcone, J. A., Carlisle, D. M., Wolock, D. M., and Meador, M. R.: GAGES: A stream gage database for evaluating natural and altered flow conditions in the conterminous United States, Ecology, 91, p. 621, 2010.

Fan, Y., Miguez-Macho, G., Weaver, C. P., Walko, R., and Robock, A.: Incorporating water table dynamics in climate modeling: 1. Water table observations and equilibrium water table simulations, J. Geophys. Res., 112, D10125, doi:10.1029/2006JD008111, 2007.

Ferguson, C. R., Sheffield, J., Wood, E. F., and Gao, H.: Quantifying uncertainty in a remote sensing based estimate of evapotranspiration over the continental United States, Int. J. Remote Sens., 31, 3821–3865, 2010.

Gao, H., Tang, Q., Ferguson, C. R., Wood, E. F., and Lettenmaier, D. P.: Estimating the water budget of major US river basins via remote sensing, Int. J. Remote Sens., 31, 3955–3978, 2010.

Garen, D. C.: Improved techniques in regression-based streamflow volume forecasting, J. Water Resour. Plan. Manag., 118, 654–670, 1992.

Gedney, N. and Cox, P. M.: The sensitivity of global climate model simulations to the representation of soil moisture heterogeneity, J. Hydrometeorol., 4, 1265–1275, 2003.

Grippa, M., Kergoat, L., Frappart, F., Araud, Q., Boone, A., de Rosnay, P., Lemoine, J. M., Gascoin, S., Balsamo, G., Ottlé, C., Decharme, B., Saux-Picart, S., and Ramillien, G.: Land water storage variability over West Africa estimated by Gravity Recovery and Climate Experiment (GRACE) and land surface models, Water Resour. Res., 47, W055489, doi:10.1029/2009WR008856, 2011.

Gulden, L. E., Rosero, E., Yang, Z. L., Rodell, M., Jackson, C. S., Niu, G. Y., Yeh, P. J. F., and Famiglietti, J.: Improving land-surface model hydrology: Is an explicit aquifer model better than a deeper soil profile?, Geophys. Res. Lett., 34, L09402, doi:10.1029/2007GL029804, 2007.

Hoerling, M., Lettenmaier, D., Cayan, D., and Udall, B.: Reconciling projections of Colorado River streamflow, Southwest Hydrol., 8, 20–21, 2009.

Huang, M., Liang, X., and Leung, L. R.: A generalized subsurface flow parameterization considering subgrid spatial variability of recharge and topography, J. Hydrometeorol., 9, 1151–1171, 2008.

Kirchner, J. W.: Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward, Water Resour. Res., 45, W02429, doi:10.1029/2008WR006912, 2009.

Koster, R. D., Mahanama, S. P. P., Livneh, B., Lettenmaier, D. P., and Reichle, R. H.: Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow, Nat. Geosci., 3,
613–616, 2010.

Krakauer, N. Y. and Temimi, M.: Stream recession curves and storage variability in small watersheds, Hydrol. Earth Syst. Sci., 15, 2377–2389, doi:10.5194/hess-15-2377-2011, 2011.

Langbein, W. B.: Chapter J: Water Budget, in: Comprehensive Survey of Sedimentation in Lake Mead, 1948–49, US Geol. Survey Professional Paper, 295, 1960.

Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models, J. Geophys. Res., 99, 14415–14428, 1994.

Liang, X., Xie, Z., and Huang, M.: A new parameterization for surface and groundwater interactions and its impact on water budgets with the variable infiltration capacity (VIC) land surface model, J. Geophys. Res., 108, 8613, doi:10.1029/2002JD003090, 2003.

Linsley, R. K., Kohler, M. A., and Paulhus, J. L. H.: Hydrology for Engineers, 3rd Edn., McGraw-Hill, New York, 508 pp., 1982.

Livneh, B., Rosenberg, E. A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K. M., Maurer, E. P., and Lettenmaier, D. P.: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous U. S.: Update and extensions, J. Climate, in review, 2013.

Mahanama, S., Livneh, B., Koster, R., Lettenmaier, D., and Reichle, R.: Soil moisture, snow, and seasonal streamflow forecasts in the United States, J. Hydrometeorol., 13, 189–203, 2012.

Maillet, E.: Essai d'hydraulique souterraine et fluviale, Libraire Sci., A. Herman, Paris, 1905.

Maurer, E. P., Wood, A. W., Adam, J. C., and Lettenmaier, D. P.: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States, J. Climate, 15, 3237–3251, 2002.

Maxwell, R. M. and Miller, N. L.: Development of a coupled land surface and groundwater model, J. Hydrometeorol., 6, 233–247, 2005.

Meyboom, P.: Estimating ground-water recharge from stream hydrographs, J. Geophys. Res., 66, 1203–1214, 1961.

Miller, J. A.: Ground Water Atlas of the United States: Introduction and National Summary (HA 730-A), US Geol. Survey, Reston, Virginia, 1999.

Mishra, V., Cherkauer, K. A., and Shukla, S.: Assessment of drought due to historic climate variability and projected future climate change in the midwestern United States, J. Hydrometeorol., 11, 46–68, 2010.

Murdock, J. N. and Calder, L.: Bank storage – Lake Powell, Bureau of Reclamation Region 4, Salt Lake City, Utah, 1969.

NRC (National Research Council): Colorado River Basin Water Management: Evaluating and Adjusting to Hydroclimatic Variability, National Academies Press, Washington, DC, 2007.

Niu, G. Y., Yang, Z. L., Dickinson, R. E., Gulden, L. E., and Su, H.: Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data, J. Geophys. Res., 112, D07103, doi:10.1029/2006JD007522, 2007.

Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., and Xia, Y.: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local scale measurements, J. Geophys. Res., 116,

D12109, doi:10.1029/2010JD015139, 2011.

Oleson, K. W., Niu, G. Y., Yang, Z. L., Lawrence, D. M., Thornton, P. E., Lawrence, P. J., Stöckli, R., Dickinson, R. E., Bonan, G. B., Levis, S., Dai, A., and Qian, T.: Improvements to the Community Land Model and their impact on the hydrological cycle, J. Geophys. Res., 113, G01021, doi:10.1029/2007JG000563, 2008.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Flanner, M. G., Kluzek, E., Lawrence, P. J., Levis, S., Swenson, S. C., Thornton, P. E., Dai, A., Decker, M., Dickinson, R., Feddema, J., Heald, C. L., Hoffman, F., Lamarque, J. F., Mahowald, N., Niu, G. Y., Qian, T., Randerson, J., Running, S., Sakaguchi, K., Slater, A., Stöckli, R., Wang, A., Yang, Z. L., Zeng, X., and Zeng, X.: Technical description of version 4.0 of the Community Land Model (CLM), NCAR Tech Note NCAR/TN-478 +STR, Natl. Cent. for Atmos. Res., Boulder, Colorado, 2010.

Quinlan, P.: Lake Mead's water level plunges as 11-year drought lingers, The New York Times, 13 August 2010, 2010.

Robson, S. G. and Banta, E. R.: Ground Water Atlas of the United States: Arizona, Colorado, New Mexico, Utah (HA 730-C), US Geol. Survey, Reston, Virginia, 1995.

Rosenberg, E. A., Wood, A. W., and Steinemann, A. C.: Statistical applications of physically based hydrologic models to seasonal streamflow forecasts, Water Resour. Res., 47, W00H14, doi:10.1029/2010WR010101, 2011.

Rutledge, A. T.: Computer programs for describing the recession of ground-water discharge and for estimating mean ground-water recharge and discharge from streamflow data – Update, Water-Resources Investigations Report 98-4148, US Geol. Survey, Reston, Virginia, 1998.

Sheffield, J., Andreadis, K. M., Wood, E. F., and Lettenmaier, D. P.: Global and continental drought in the second half of the twentieth century: Severity-area-duration analysis and temporal variability of large-scale events, J. Climate, 22, 1962–1981, 2009.

Shukla, S. and Lettenmaier, D. P.: Seasonal hydrologic prediction in the United States: understanding the role of initial hydrologic conditions and seasonal climate forecast skill, Hydrol. Earth Syst. Sci., 15, 3529–3538, doi:10.5194/hess-15-3529-2011, 2011.

Speers, D. D., Rockwood, D. M., and Ashton, G. D.: Chapter 7: Snow and snowmelt, in: Hydrology Handbook, A. S. C. E. Manuals and Reports on Engineering Practice No. 28, 2nd Edn., Am. Soc. Civil Eng., New York, 437–476, 1996.

Stern, P. C. and Easterling, W. E. (Eds.): Making Climate Forecasts Matter, National Research Council Report, National Academies Press, Washington, DC, 1999.

Strassberg, G. B., Scanlon, R., and Chambers, D.: Evaluation of groundwater storage monitoring with the GRACE satellite: Case study of the High Plains aquifer, central United States, Water Resour. Res., 45, W05410, doi:10.1029/2008WR006892, 2009.

Swenson, S. C. and Wahr, J.: Post-processing removal of correlated errors in GRACE data, Geophys. Res. Lett., 33, L08402, doi:10.1029/2005GL025285, 2006.

Syed, T. H., Famiglietti, J. S., Rodell, M., Chen, J., and Wilson, C. R.: Analysis of terrestrial water storage changes from GRACE and GLDAS, Water Resour. Res., 44, W02433, doi:10.1029/2006WR005779, 2008.

Tang, Q., Peterson, S., Cuenca, R. H., Hagimoto, Y., and Lettenmaier, D. P.: Satellite-based near-real-time estimation of irrigated crop water consumption, J. Geophys. Res., 114, D05114, doi:10.1029/2008JD010854, 2009.

Troccoli, A., Harrison, M., Anderson, D. L. T., and Mason, S. J. (Eds.): Seasonal Climate: Forecasting and Managing Risk, NATO Science Series IV: Earth and Environmental Sciences, Springer Academic Publishers, 499 pp., 2008.

USDOI (US Department of Interior): Colorado River interim surplus criteria, Final Environmental Impact Statement, Volume 1, Bureau of Reclamation, 2000.

Vogel, R. M. and Kroll, C. N.: Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics, Water Resour. Res., 28, 2451–2458, doi:10.1029/92WR01007, 1992.

Whitehead, R. L.: Ground Water Atlas of the United States: Montana, North Dakota, South Dakota, Wyoming (HA 730-I), US Geol. Survey, Reston, Virginia, 1996.

Wood, A. W. and Lettenmaier, D. P.: A test bed for new seasonal hydrologic forecasting approaches in the western United States, B. Am. Meteorol. Soc., 87, 1699–1712, 2006.

Wood, A. W. and Lettenmaier, D. P.: An ensemble approach for attribution of hydrologic prediction uncertainty, Geophys. Res. Lett., 35, L14401, doi:10.1029/2008GL034648, 2008.

Yeh, P. J. F. and Eltahir, E. A. B.: Representation of water table dynamics in a land surface scheme, Part I: Model development, J. Climate, 18, 1861–1880, 2005.

Zampieri, M., Serpetzoglou, E., Anagnostou, E. N., Nikolopoulos, E. I., and Papadopoulos, A.: Improving the representation of river–groundwater interactions in land surface modeling at the regional scale: Observational evidence and parameterization applied in the Community Land Model, J. Hydrol., 420–421, 72–86, 2012.

Zhang, K., Kimball, J. S., Nemani, R. R., and Running, S. W.: A continuous satellite derived global record of land surface evapotranspiration from 1983–2006, Water Resour. Res., 46, W09522, doi:10.1029/2009WR008800, 2010.