Journal of Hydrology 529 (2015) 535–547

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

A water balance approach for reconstructing streamflow using tree-ring proxy records

Laurel Saito [a,*], Franco Biondi [b], Rajan Devkota [c], Jasmine Vittori [d,1], Jose D. Salas [e]

[a] Graduate Program of Hydrologic Sciences, Department of Natural Resources and Environmental Science, University of Nevada Reno, 1664 N. Virginia Street, Mail Stop 0186, Reno, NV 89557, USA
[b] Dendrolab and Graduate Program of Hydrologic Sciences, University of Nevada Reno, 1664 N. Virginia Street, Mail Stop 0154, Reno, NV 89557, USA
[c] Graduate Program of Hydrologic Sciences, University of Nevada Reno, 1664 N. Virginia Street, Mail Stop 0175, Reno, NV 89557, USA
[d] Graduate Program of Environmental Sciences, University of Nevada Reno, 1664 N. Virginia Street, Mail Stop 0186, Reno, NV 89557, USA
[e] Department of Civil and Environmental Engineering, Colorado State University, Fort Collins, CO 80523, USA

ARTICLE INFO

Article history:
Available online 18 November 2014

Keywords:
Modeling
Dendrochronology
Paleohydrology
Walker River
Watersheds
Runoff

SUMMARY

Tree-ring data have been used to augment limited instrumental records of climate and provide a longer view of past variability, thus improving assessments of future scenarios. For streamflow reconstructions, traditional regression-based approaches cannot examine factors that may alter streamflow independently of climate, such as changes in land use or land cover. In this study, seasonal water balance models were used as a mechanistic approach to reconstruct streamflow with proxy inputs of precipitation and air temperature. We examined a Thornthwaite water balance model modified to have seasonal components and a simple water balance model with a snow component. These two models were calibrated with a shuffled complex evolution approach using PRISM and proxy seasonal temperature and precipitation to reconstruct streamflow for the upper reaches of the West Walker River basin at Coleville, CA. Overall, the modified Thornthwaite model performed best during calibration, with R^2 values of 0.96 and 0.80 using PRISM and proxy inputs, respectively. The modified Thornthwaite model was then used to reconstruct streamflow during AD 1500–1980 for the West Walker River basin. The reconstruction included similar wet and dry episodes as other regression-based records for the Great Basin, and provided estimates of actual evapotranspiration and of April 1 snow water equivalence. Given its limited input requirements, this approach is suitable in areas where sparse instrumental data are available to improve proxy-based streamflow reconstructions and to explore non-climatic reasons for streamflow variability during the reconstruction period.

1. Introduction

Long-term records of water resources are important for sustainable water management, especially in the Western United States where population growth is high and climate in many regions is semi-arid with frequent drought and low rainfall (Anderson and Woosley, 2005). In addition, limited water supplies in the Western United States are anticipated to further decrease due to increased demand and predicted climate change (Cayan et al., 2010; Milly et al., 2005; Seager et al., 2007). Instrumental records of precipitation, temperature, and surface-water flow in this part of the world are usually limited to the last century, but long-term estimates of streamflow variability are critical for managing the impacts of floods and droughts (Stewart et al., 2004; Wilhite, 2000). Among proxies used for extending climate records and providing a paleo-perspective on recent changes, tree-ring records, having annual to seasonal resolution, allow for numerical calibration with instrumental records (Cook and Kairiukstis, 1990; Fritts, 1976).

Dendroclimatic studies have generated multi-century long series of temperature, precipitation, streamflow, and drought all over the world (Frank et al., 2005; Tingley and Huybers, 2010), including the Western United States (Meko and Graybill, 1995; Woodhouse and Lukas, 2006). These reconstructions are usually based on linear regression techniques, but when applied to streamflow, such empirical approaches cannot directly test the influence of watershed factors such as changes in land use or soil type that

* Corresponding author at: University of Nevada Reno, 1664 N. Virginia Street, Mail Stop 0186, Reno, NV 89557, USA. Tel.: +1 775 784 1921.
E-mail addresses: lsaito@cabnr.unr.edu (L. Saito), fbiondi@unr.edu (F. Biondi), rajandevkota@gmail.com (R. Devkota), jasmine.vittori@gmail.com (J. Vittori), jsalas@engr.colostate.edu (J.D. Salas).
[1] Present address: Stantec, 595 Double Eagle Ct., Suite 2000, Reno, NV 89521, USA.

http://dx.doi.org/10.1016/j.jhydrol.2014.11.022
0022-1694/

can alter streamflow independently of climate. A mechanistic approach could therefore be useful in these cases because non-climatic factors could be inserted in model runs to test the sensitivity of the reconstructed streamflow and obtain more realistic uncertainty estimates (Saito et al., 2008).

Mechanistic watershed models are fundamental to water resources assessment, development, and management (Singh and Woolhiser, 2002). These models have been used to analyze the quantity and quality of streamflow, reservoir system operations, surface water and groundwater conjunctive use, and a range of other water management activities (Wurbs, 1998; Xu, 2002). Most watershed models generally require detailed hydrological and meteorological inputs. Data requirements often include minimum and maximum air temperatures, short-wave radiation (especially for snowmelt calculations), pan evaporation, precipitation, dew-point temperature, relative humidity, wind speed, and watershed data such as topography, soils, and vegetation features (Singh and Woolhiser, 2002). Most of these data are not available when streamflow is reconstructed using tree-ring records over several centuries, which is one reason why regression-based approaches have been preferred.

Among mechanistic watershed modeling approaches, a water balance model can be considered as a system of equations designed to represent components of the hydrologic cycle. Water balance models were first developed in the 1960s, and have been used over time for various water management issues (Alley, 1984; Xu, 2002). For instance, these models have been applied for determining seasonal patterns of irrigation demand, soil moisture stress, and prediction of streamflow. Water balance models keep track of water input and output by accounting for precipitation and snowmelt, evapotranspiration, streamflow, runoff, and groundwater (Alley, 1984), and can range from quite simple to very complex, depending on study objectives.

Some recent studies have examined the use of water balance models with proxy data to model hydrologic processes for surface runoff, evapotranspiration (ET), groundwater flow, base flow, and snowmelt using as input either precipitation alone (Saito et al., 2008) or both precipitation and temperature (Gray and McCabe, 2010; Solander et al., 2010). Saito et al. (2008) used a water balance model adapted from Fiering (1967), Salas and Obeysekera (1992), and Vogel and Sankarasubramanian (2003) to simulate streamflows in the upper West Walker River watershed in California on a water year (WY; October 1 through September 30) timescale. Total WY precipitation obtained from the Parameter-elevation Regressions on Independent Slopes Model (PRISM; Daly et al., 1994) was the only input, and the model represented observed streamflow quite well ($R^2 = 0.87$, $n = 63$ years). Since this model did not include a snow component, it only had four parameters. Solander et al. (2010) modified the model used by Saito et al. (2008) to operate on a seasonal timescale, and used a temperature-index approach for addressing snowfall and snowmelt. PRISM-derived cold season precipitation and air temperature were used as input, and the model was calibrated to the upper Meadow Valley Wash watershed in southeastern Nevada ($R^2 = 0.81$; $n = 18$ years). However, this particular basin had the added complications of a downstream dam and upstream diversions for seasonal irrigation, which both influence streamflow during part of the year. Thus, this model was only used for the cold season, the part of the year with natural flow regimes because diversions were not supposed to occur.

Gray and McCabe (2010) utilized the monthly Thornthwaite water balance model (McCabe and Markstrom, 2007) to generate annual streamflows for the Upper Yellowstone drainage and simulate impacts of future climate scenarios on regional hydrology. Model calibration resulted in a correlation coefficient of 0.71 for the period 1911–1995 using PRISM precipitation and air temperature data as input. Proxy precipitation data were then obtained at annual resolution by calibrating tree-ring chronologies against PRISM precipitation between 1895 and 2004. When PRISM temperature and proxy precipitation were used as input, model calibration resulted in a correlation coefficient of 0.56.

Because proxy records of precipitation and temperature are available at annual and seasonal time scales at best (Bradley, 1999), our objective was to improve upon a seasonal water balance model that could use proxy precipitation and temperature as input to estimate past streamflow over long time periods. We focused on snowpack dynamics because in mountainous regions, snowfall significantly affects watershed hydrology by storing water in the form of snow and releasing it either in part or in full as snowmelt in the spring, depending on climatic conditions (Pavelsky et al., 2011; Stewart, 2009). In the Western United States, approximately 50–80% of the total annual water budget comes from snowmelt (Hanson, 2001; Serreze et al., 2001; Williams and Tarboton, 1999).

Our models were applied to the upper West Walker River Basin, where water issues have generated heated debate in recent times (Remick, 2003) and millennia-long tree-ring records are available (Biondi et al., 2008). The main challenge was to obtain reasonable streamflow estimates while using a minimal number of model parameters, and still allowing for the inclusion of non-climatic factors in model simulations to better constrain the uncertainty of the reconstructed streamflow. Models were calibrated against stream gauge data for the West Walker River basin at the United States Geological Survey (USGS) station of Coleville, California to determine the most appropriate approach for reconstructing streamflow with seasonal proxy inputs.

2. Materials and methods

2.1. Site description

The upper West Walker River watershed upstream of the Coleville station has an area of 490 km^2 and is located near the boundary between California and Nevada (Fig. 1). Streamflow gauge data (USGS gauge 10296000, 38.38°N, 119.45°W, elevation 2009 m; http://waterdata.usgs.gov/nwis/nwisman/?site_no=10296000&agency_cd=USGS) are available since 1938. The West Walker River originates in the Eastern Sierra Nevada and flows north–northeast before converging with the East Walker River in Mason Valley just south of Yerington, Nevada. Upon exiting Mason Valley, the river turns southward before terminating in Walker Lake, a remnant of prehistoric Lake Lahontan that encompassed most of the Great Basin during the last ice age (Horton, 1996). The watershed elevation ranges from about 2000 m to more than 3700 m. Snow water equivalent (SWE) data were available from five sites: Leavitt Meadows, Sonora Pass, Virginia Lakes, Virginia Lakes Ridge, and Willow Flat (California Data Exchange; http://cdec.water.ca.gov/cgi-progs/snowQuery).

There is no meteorological station in the upper West Walker River watershed and the nearest one is in Bridgeport, CA (38.25°N, 119.23°W, elevation 1972 m; Fig. 1). Data from this meteorological station go back to 1948 and show an average calendar year precipitation of about 230 mm. Average calendar year snowfall from the nearest SNOTEL site in the upper West Walker River watershed (Sonora Pass, 38.32°N, 119.6°W, elevation 2682 m; in operation since 1984) is approximately 1100 mm. From January 1895 to December 2001, average annual WY precipitation in 12 PRISM cells that cover the watershed is about 1200 mm, and average monthly PRISM temperatures range from −10 °C in February to 23 °C in July.

The geology of the Walker River Basin centers around the modern Sierra Nevada batholiths, which formed around 225 to 65 million years ago (Hill, 1975). Natural processes eroded the overlying

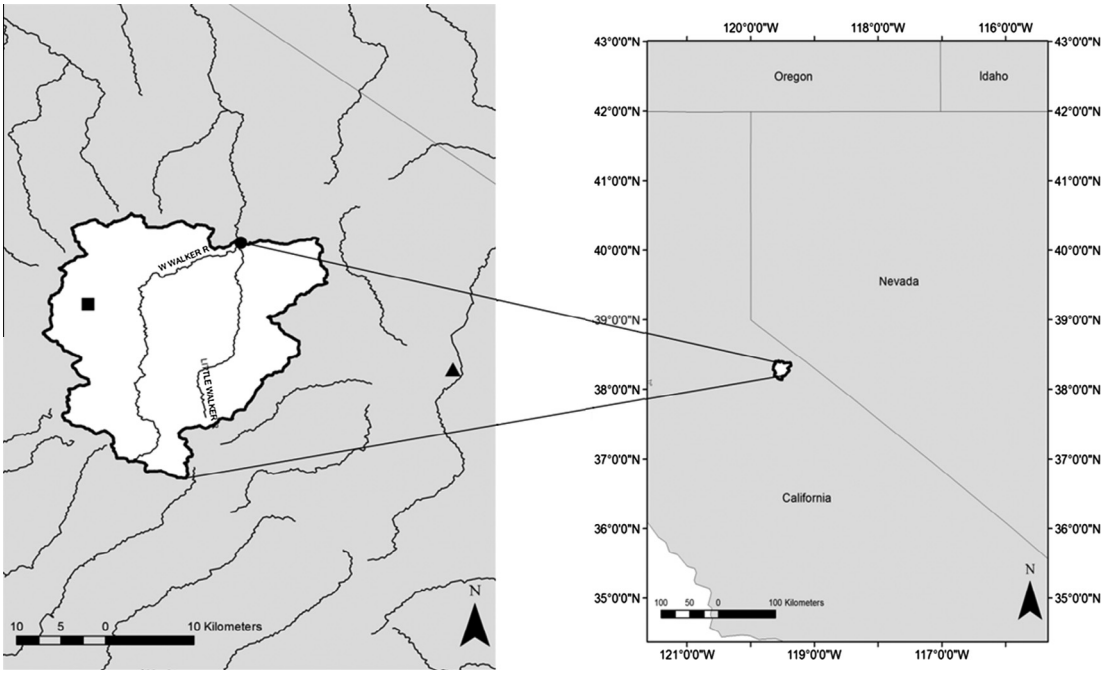

Fig. 1. Left: Map of the upper West Walker River watershed, including the USGS streamflow station at Coleville (solid circle), streams (lines), watershed boundary (heavy black line), SNOTEL site (square), and Bridgeport meteorological station (triangle). Right: Watershed location in the Western United States.

external rock, exposing the batholiths in the region as an area of low, gently rolling granitic hills. Around 20 million years ago volcanic eruptions blanketed the area with lava flows and volcanic material (Reed, 1933). As eruptions increased in frequency, the eastern edge of the modern Sierra Nevada mountain range began to rise and tilt along fault lines. These processes of uplift and erosion continue to shape and wear down the Sierra Nevada mountain range (Hill, 1975). The eastern side of the range is considerably more tilted and is less weathered due in part to the rain shadow on this side of the range (Reed, 1933). Human settlement within the watershed is minimal and has not changed much over the period of instrumental streamflow records. Lightning-caused wildfires are a natural occurrence in this landscape, and 14 wildfires were documented in the upper West Walker River Basin between WY 1961 and 2006 (Vittori, 2011). All wildfires occurred during the dry season, with seven fires recorded in 1991 at various locations in the watershed.

The vegetation in the watershed is directly related to elevation and orographic precipitation, both of which decline from west to east. Higher elevations have mixed conifer stands, with western white pine (*Pinus monticola*), red fir (*Abies magnifica*), Jeffrey pine (*Pinus jeffreyi*) at the uppermost elevations, and white fir (*Abies concolor*) and western juniper (*Juniperus occidentalis*) at mid-elevations (Kattelmann, 2012). Brushfields also are present at higher elevations with buckbrush (*Ceonothus velutinus*) and chokecherry (*Prunus emarinatus*), whereas at lower elevations sage (*Artemesia tridentata*), bitterbrush (*Purshia tridentata*), mountain mahogany (*Cercocarpus ledifolius*), and snowberry (*Symphoricarpos albus*) dominate brush communities (Kattelmann, 2012).

2.2. Model approaches

2.2.1. Thornthwaite seasonal model (Model A)

The Thornthwaite seasonal model (Fig. 2a) is a modification of the Thornthwaite monthly water balance model that allocates water among various components of the hydrologic system (McCabe and Wolock, 1999; McCabe and Markstrom, 2007; Thornthwaite, 1948). Inputs to the model are seasonal temperature (T, °C), seasonal total precipitation (P, mm), and latitude (°) of the location used for computing day length and then potential evapotranspiration (PET). The Thornthwaite model's monthly day length factor was converted to a seasonal day length by averaging the monthly day length factors for each season (October to March for cold season, and April to September for warm season) and dividing the average value by the number of days in each season (182 and 183 for the cold and warm seasons, respectively).

The model estimates the amount of seasonal precipitation (P_{total}) that is rain (P_{rain}) or snow (P_{snow}) in mm for each timestep. When seasonal temperature (T) is below a specified threshold (T_{snow}, °C), all precipitation is considered to be snow. If temperature is greater than an additional threshold (T_{rain}, °C), then all precipitation is considered to be rain. When the seasonal air temperature is between T_{snow} and T_{rain}, the amount of precipitation that is snow decreases linearly from 100% to 0% of the total precipitation according to Eq. (1):

$$P_{snow} = P_{total} \times \left[\frac{T_{rain} - T}{T_{rain} - T_{snow}}\right] \quad (1)$$

P_{rain} is then computed as:

$$P_{rain} = P_{total} - P_{snow} \quad (2)$$

Direct runoff (DRO, mm) is runoff from impervious surfaces, or runoff resulting from infiltration excess overflow. DRO is calculated as:

$$DRO = P_{rain} \times drofrac \quad (3)$$

where $drofrac$ is a direct runoff coefficient that represents the fraction of P_{rain} that becomes runoff. DRO is subtracted from P_{rain} to compute the amount of remaining precipitation (P_{remain}):

$$P_{remain} = P_{rain} - DRO \quad (4)$$

Snow storage ($snoster$, mm) is subjected to melt if conditions are such that melting can occur. Thus, snow, rain and snowmelt can occur in the same season. The fraction of $snoster$ that melts

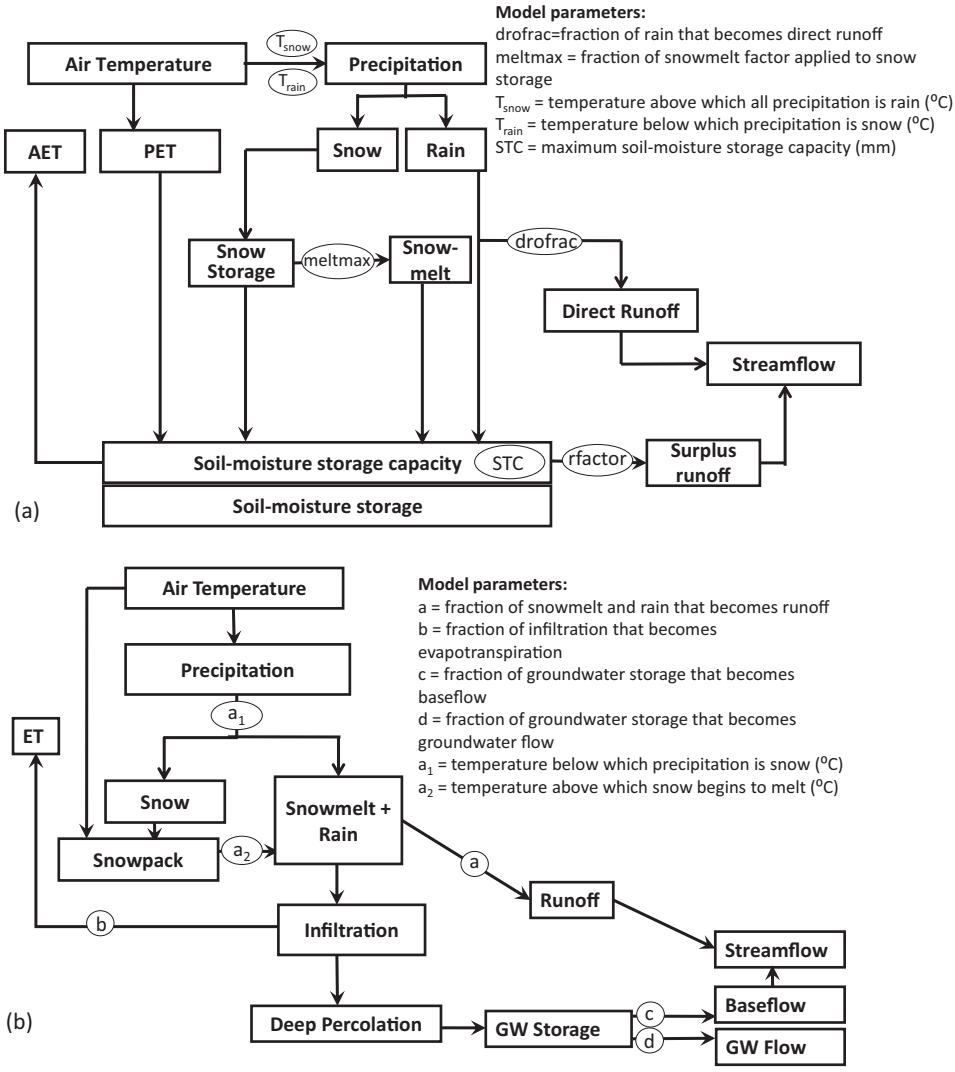

Fig. 2. Schematics of (a) the modified Thornthwaite seasonal model (Model A) and (b) the simple water balance model with WASMOD snow component (Model B).

in a season (snowmelt fraction, *SMF*) is computed from seasonal average temperature (*T*) and a maximum melt rate (*meltmax*). Maximum melt rate (*meltmax*) is one of the calibrated parameters (McCabe and Wolock, 1999). The fraction of snow storage that melts (*SMF*) in a season is computed as:

$$SMF = \left[\frac{T - T_{snow}}{T_{rain} - T_{snow}}\right] \times meltmax \quad (5)$$

If the computed *SMF* is greater than *meltmax*, then *SMF* is set to *meltmax*. The amount of snow that is melted in a season (*SM*) in mm of snow water equivalent is computed as:

$$SM = snoster * SMF \quad (6)$$

SM is added to P_{remain} to compute the total liquid water input (P_{liquid}) to the soil. Additional equations for the model are included in Appendix A. Calibration parameters for the modified Thornthwaite model are T_{rain}, T_{snow}, *drofrac*, *meltmax*, *rfactor*, *STC*, *GS*, and *SW*, where *STC* is the soil moisture storage capacity (cm), *GS* is initial groundwater storage (cm), and *SW* is initial surface water (cm).

2.2.1. Simple water balance model with WASMOD (Model B)

The simple water balance model was modified from that used by Saito et al. (2008), which required an initial boundary condition of starting groundwater storage (*GS*) plus four water balance parameters that did not change with time. These parameters were *a* = fraction of snowmelt and rain that becomes surface runoff; *b* = fraction of infiltrated water that evapotranspires; *c* = fraction of groundwater storage that becomes base flow; and *d* = fraction of groundwater storage that becomes groundwater flow. To convert this annual model to a seasonal one, it was necessary to add a snow component that was derived from the Water and Snow balance MODeling system (WASMOD; Fig. 2b; Xu, 2002). To determine the amount of snow that melts, two threshold temperature parameters, a_1 (°C) and a_2 (°C), were used, with $a_1 \geqslant a_2$. Snowmelt begins when air temperature is higher than a_2; snowfall starts to become rainfall when air temperature is higher than a_1. These threshold temperatures were calibration parameters and their acceptable ranges were 0–3.89 °C for a_1 and 0–a_1 °C for a_2 (Xu, 2002).

The snow calculation procedure in WASMOD begins by calculating the average seasonal snowfall component Sn_t (cm) of precipitation P_t (cm) as a fraction of seasonal P_t that depends on the average seasonal temperature *T* (°C):

$$Sn_t = P_t\left[1 - e^{\left[\frac{(T-a_1)}{(a_1-a_2)}\right]^2}\right] \quad (7)$$

Seasonal rainfall, R_t (cm), is therefore:

$$R_t = P_t - Sn_t \quad (8)$$

Snowmelt during each season t is calculated as a function of the average seasonal temperature T, and snowpack Sp_{t-1} at the beginning of season t as:

$$M_t = Sp_{t-1}\left[1 - e^{\left[\frac{(T-a_2)}{(a_1-a_2)}\right]^2}\right] \quad (9)$$

The snowpack balance equation is:

$$Sp_t = Sp_{t-1} + Sn_t - M_t \quad (10)$$

The remaining snowpack Sp_t is added to the snowfall for the next season. The simple water balance model with WASMOD snow component has seven calibration parameters: a_1, a_2, a, b, c, d, and GS. Additional equations for the model are included in Appendix A.

2.2.3. Model inputs and setup

For Model A, upper and lower bounds of parameter values were selected according to available literature for the Thornthwaite model (Table 1; Gray and McCabe, 2010; McCabe and Markstrom, 2007; Rango and Martinec, 1995; Thornthwaite and Mather, 1955; Wolock and McCabe, 1999). Parameter a of Model B represents the fraction of snowmelt and precipitation that becomes runoff, and theoretically the upper bound for the parameter is equal to one. Because the dominant soil group in the region is Hydrologic Soil Group D (Natural Resources Conservation Service, 2009), which has a very slow rate of water transmission, parameter a is expected to be high so its upper bound was set at 0.9. The upper boundaries for b, c, and d for Model B were set to 1.0 because ET, base flow, and groundwater (GW) flow could not exceed available infiltrated water or GW storage (Saito et al., 2008). Lower boundary conditions for a, b, c, and d were set to 0.0. The upper boundary of the sum of parameters c and d was fixed at 1 so that the base flow and GW flow could not exceed the available groundwater storage.

2.2.3.1. Input data. The seasonal water balance models used seasonal temperature and precipitation calculated from monthly PRISM precipitation and maximum and minimum temperature at 2.5 arc-minute resolution. PRISM grid cells included in the watershed area provided data for WY 1940–2011. Seasons were defined as cold–wet (October to March) and warm–dry (April to September) based on precipitation data, with cold–wet precipitation accounting for about 80% of the total WY precipitation on average. Monthly maximum and minimum PRISM air temperatures were averaged over each season to get maximum and minimum seasonal temperatures. In addition, monthly maximum and minimum temperatures were averaged to get monthly average temperatures which were then averaged by season to get average seasonal temperatures.

In the modified Thornthwaite model (Model A), data for the average mean seasonal temperature (T) in Eq. (1) were the average maximum seasonal temperature (T_{max}; Model A-1), the average minimum seasonal temperature (T_{min}; Model A-2), or the average mean seasonal temperature (T_{avg}; Model A-3). For the simple water balance model with WASMOD snow component (Model B), data for the average seasonal temperature (T) in Eqs. (7) and (9) were the average maximum seasonal temperature (T_{max}; Model B-1), average minimum seasonal temperature (T_{min}; Model B-2), or the average mean seasonal temperature (T_{avg}; Model B-3).

WY proxy precipitation in cm (Fig. 3a) was calculated by regressing PRISM precipitation for WY 1896–2001 against precipitation values in standard deviation units (SDUs) from Biondi et al. (2008). Mean WY proxy temperature in °C (Fig. 3b) was calculated by regressing PRISM temperature for WY 1896–1980 against temperature anomalies from Wahl and Smerdon (2012) for their grid cell (37.5°N, 117.5°W) nearest to the West Walker River at Coleville, CA. The temperature and precipitation regressions between PRISM and proxy records resulted in R^2 values of 0.27 and 0.35, respectively.

Table 1
Boundary values and constraints for Model A (modified Thornthwaite model) and Model B (simple water balance model with WASMOD snow component).

Parameter (units)	Model A	Model B
T_{rain} (°C)	0 – 5	
T_{snow} (°C)	-1 - (-10)	
drofrac	0.1 - 1	
meltmax	0.01 - 1	
STC (cm)	100 - 250	
rfactor	0 – 1	
a_1(°C)		0 - 3.89
a_2(°C)		0 - a_1
a		0 - 0.9
b		0 - 1
c		0 - 1
d		0 - 1
Initial groundwater storage (GS) (cm)	0 - 200	0 - 200
$c + d$		≤ 1
Initial surface water (cm)	0 - 100	

Shading indicates parameters not used in a particular model.

Precipitation and temperature records derived from tree rings were on an annual scale, hence disaggregation into a seasonal time scale was necessary for use in our seasonal water balance models. Using historical (WY 1940–2011) PRISM data, we found that the ratio of warm–dry season to WY precipitation decreases as WY precipitation increases. Thus, annual values of proxy precipitation from WY 1500 to 1980 AD were divided into quartiles, and modern ratios (Table 2) were used to disaggregate proxy WY precipitation into warm season precipitation (Gray and McCabe, 2010). Cold–wet season precipitation values were calculated by subtracting the warm–dry season precipitation from the WY value each year. Linear regression coefficients estimated between PRISM average WY temperatures and PRISM warm–dry (R^2 = 0.42) or cold–wet (R^2 = 0.47) seasonal temperatures for WY 1896–1980 were used to estimate warm–dry and cold–wet season temperatures from the proxy annual temperatures for WY 1500 to 1980.

2.2.4. Model calibration, validation, and verification

The models were tested to determine appropriate parameters to use for proxy reconstructions by applying the shuffled complex evolution (SCE) method, which is a global optimization strategy designed and developed by Duan et al. (1992). Various researchers have investigated the use of the SCE method for calibration of hydrological and water balance, soil erosion, subsurface hydrology, remote sensing, and land surface models and have found it to be efficient, robust, and consistent in comparison to other optimization strategies (Eckhardt et al., 2005; Gupta et al., 1999; Thyer et al., 1999; Vanhaute et al., 2012; Xiao et al., 2009).

The SCE algorithm in the first step (zero-loop) generates a "population" of points by random sampling throughout the feasible parameter space for the parameters to be optimized. A criterion value (objective function) is calculated for each parameter set and the sets are ranked in order of increasing criterion value. The population is then partitioned into a number of "complexes," each consisting of $2n + 1$ parameter sets, where n is the number of parameters to be optimized. Each complex then evolves independently following a Competitive Complex Evolution (CCE) algorithm which is based on a simplex downhill search scheme (Nelder and Mead, 1965). The parameter sets in the evolved complexes are next combined into a single sample population and ranking is done based on the criterion value. The entire population is then shuffled or re-partitioned by re-assigning the parameter sets into new

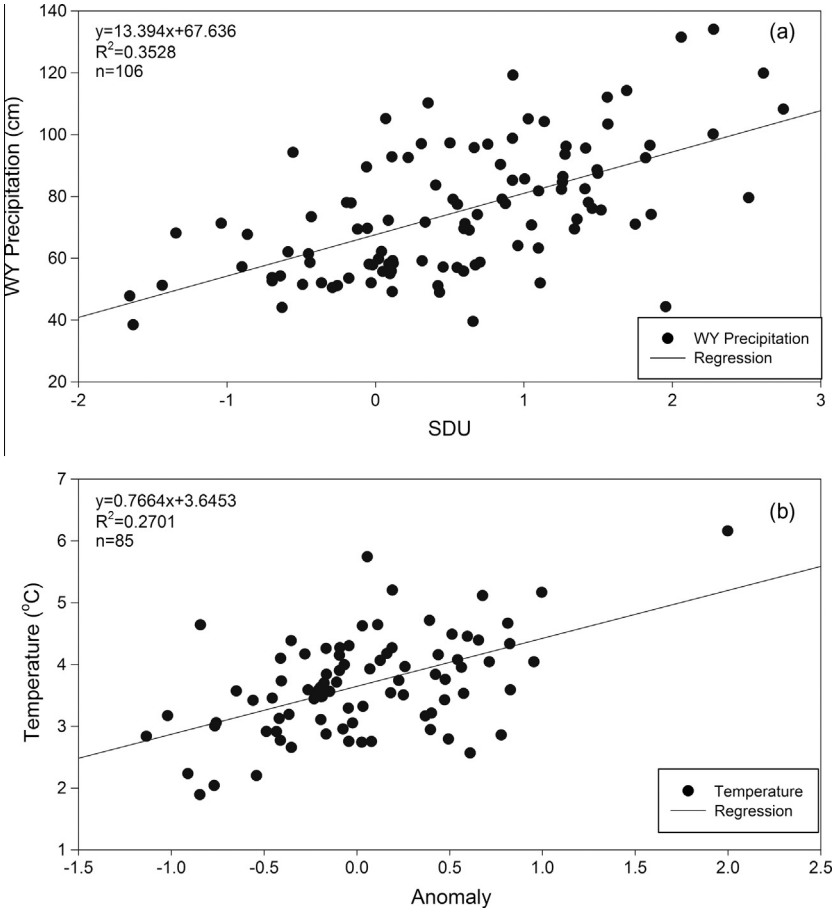

Fig. 3. Linear regressions of (a) water year (WY) 1896–2001 PRISM precipitation against tree-ring reconstructed precipitation (standard deviation units, SDU) from Biondi et al. (2008); and (b) PRISM temperature for WY 1896–1980 against tree-ring reconstructed temperature anomalies from Wahl and Smerdon (2012).

Table 2
Ratio of PRISM warm–dry season precipitation to PRISM WY precipitation according to precipitation quartiles for WY 1940–2011.

Quartile	Ratio
1st	0.271
2nd	0.211
3rd	0.185
4th	0.180

complexes formed so that information gained by the previous complexes is shared. The evolution and shuffling steps continue until pre-specified convergence criteria are reached.

The combination of deterministic and probabilistic approaches requires some algorithmic parameters that control these components. They are m, the number of points in a complex; q, the number of points in a subcomplex; p, the number of complexes; $pmin$, the minimum number of complexes required in the population; α, the number of consecutive offspring generated by each sub-complex; and β, the number of evolution steps taken by each complex. Duan et al. (1994) and Sorooshian et al. (1993) recommended some default values and equations to compute these parameters: $m = 2n + 1$; $q = n + 1$; $\alpha = 1$; $\beta = m$; and $p = 20$.

The SCE method was used to calibrate the water balance models with PRISM precipitation and temperature inputs. These models used default parameter values for the SCE algorithms. Model calibration and validation with PRISM inputs followed the classical split-sample test as defined by Klemes (1986). The entire record (WY 1940–2011) was divided into two time periods: WY 1940–1978 and WY 1979–2011. The first time period was used for calibration and the second for validation, and then the second time period was used for calibration and the first for validation. Parameter estimates obtained in the calibration period were used in the validation run. SCE terminated when the objective function was reduced by one percent or less over five successive iterations (i.e. evolution loops), or if a maximum of 50,000 model runs had been carried out.

The Thornthwaite seasonal water balance model (Model A) and the simple water balance model with WASMOD snow component (Model B) were programmed in MATLAB and the SCE algorithm was implemented to calibrate eight parameters for Model A and seven parameters for Model B using upper and lower bounds of these parameters in Table 1.The objective function was to minimize the root mean-squared error (RMSE) between actual streamflow recorded at the USGS streamflow gage in Coleville, CA (i.e., mean daily streamflow for each seasonal period) and simulated streamflow (Eq. (11)):

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(Q_t^{sim} - Q_t^{obs}\right)^2} \quad (11)$$

where n is the number of observations (in this case the total number of seasons), and Q_t^{sim} and Q_t^{obs} are respectively defined as the simulated and observed streamflow values (in cm per season) at timestep t. In addition to RMSE, R^2, percent bias, relative RMSE (RRMSE; Eq. (12)), and Nash Sutcliffe Efficiency (NSE; Eq. (13)) were calculated to evaluate model performance:

$$RRMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\frac{(Q_t^{sim} - Q_t^{obs})}{Q_t^{obs}}\right)^2} \quad (12)$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Q_t^{sim} - Q_t^{obs})^2}{\sum_{i=1}^{n}(Q_t^{obs} - \bar{Q}_t^{obs})^2} \quad (13)$$

where $\bar{Q}^{obs}$ is defined as the average of observed streamflow. These model performance measures were also calculated for each season (warm–dry and cold–wet), and all metrics were evaluated to select the best model.

SCE was also used to calibrate Model A and Model B with proxy seasonal input data for WY 1940–1980. Because the period of overlap between proxy data and instrumental streamflows was so limited, only calibration was performed (i.e., no validation, reverse calibration or reverse validation).

2.2.5. Sensitivity analysis and streamflow reconstruction

Because one of the values of the new approach could be the ability to simulate "what if" scenarios regarding changes in the watershed, sensitivity analysis of model parameters was performed to provide insight into the robustness of parameter values and their uncertainties by demonstrating the influence of the values of a parameter on simulation results (Bardossy and Singh, 2008). Parameters that have larger influences on simulated results with small changes in their values are more sensitive than those have little influence when changed (Brath et al., 2004; Sieber and Uhlenbrook, 2005).

We tested the sensitivity of model parameters by changing their values in 10% increments within their feasible ranges while keeping other parameters constant at the SCE calibrated values with PRISM inputs. The resulting RMSEs were calculated to identify parameters with large influences on simulated results (Brath et al., 2004; Sieber and Uhlenbrook, 2005).

Finally, we used the calibrated parameters for the best performing model to reconstruct streamflows with proxy precipitation and temperature for the period 1500–1980. The resulting time series of streamflows was evaluated using the methods of Biondi et al. (2008) to quantify and rank positive and negative episodes.

3. Results

Calibration results using SCE with PRISM inputs (Table 3) showed that the Thornthwaite seasonal water balance model with average minimum temperatures (Model A-2) and the simple water balance model with WASMOD snow component and average seasonal temperatures (Model B-3) performed better than other versions of the same models. Similar results were obtained for the validation, reverse calibration, and reverse validation runs using SCE with PRISM inputs (Table 3). The best overall calibration performance metrics (i.e., metrics for results for both seasons together) were found for the Thornthwaite seasonal model with minimum average seasonal temperatures (Model A-2) and included an overall R^2 of 0.96, RMSE of 4.57 cm, Bias of −3.73%, NSE of 0.95, and RRMSE of 0.29, indicating that model simulations compared well with observed values (Fig. 4a). However, model performance worsened during the reverse calibration period (Table 3). The simple water balance model (Model B-3) performed well with observed values during the calibration and reverse calibration runs, with R^2 of 0.94 (Table 3; Fig. 4b). Validation and reverse validation statistics indicated that Model B-3 performed better outside of the calibration period compared to Model A-2 (Table 3). Most performance metrics for all models were worse during the cold–wet season compared to the warm–dry season. For example, R^2 values were 0.75–0.90 in the warm–dry season, and 0.18–0.52 in the cold–wet one for calibration and reverse calibration.

SCE calibrations for Model A and Model B using the tree-ring derived cold–wet and warm–dry seasonal precipitation and temperature for WY 1940–1980 (Table 4) indicated the modified Thornthwaite model (Model A) performed better than Model B. The best calibration performance metrics for this model were an overall R^2 of 0.80, RMSE of 10.71 cm, bias of −19.33%, NSE of 0.75, and RRMSE of 0.65. Cold–wet seasonal performance metrics were usually better than warm–dry seasonal metrics, which was also the case in the SCE calibration using PRISM inputs. Overall model performance worsened using tree-ring derived precipitation and temperature as compared to PRISM data, but remained better than for the Gray and McCabe (2010) watershed model with tree-ring inputs.

Table 3
Performance metrics of all models from the Shuffled Complex Evolution (SCE) approach for calibration period (cal.; WY 1940–1978). Also shown are performance metrics of the best-performing models for validation (val; WY 1979–2011), reverse calibration (rev. cal.; WY 1979–2011) and reverse validation (rev. val.; WY 1940–1978) with PRISM input data.

	Model A						Model B					
Performance metrics (units)	A-1 cal.	A-2 cal.	A-3 cal.	A-2 val.	A-2 rev. cal.	A-2 rev. val.	B-1 cal.	B-2 cal.	B-3 cal.	B-3 val.	B-3 rev. cal.	B-3 rev. val.
R^2 overall	0.00	***0.96***	0.94	0.83	0.89	0.93	0.91	0.91	**0.94**	0.88	0.91	0.93
R^2 warm	0.08	**0.88**	0.87	0.65	0.75	0.84	0.80	0.82	***0.90***	0.76	0.81	0.87
R^2 cold	0.00	**0.44**	0.43	0.15	0.18	0.02	0.00	0.00	***0.52***	0.04	0.23	0.23
RMSE overall (cm)	27.82	***4.57***	9.34	10.81	8.70	7.16	7.04	6.97	**5.27**	8.40	7.39	5.70
RMSE warm (cm)	38.47	***5.66***	11.22	14.06	11.55	9.31	7.70	7.35	**6.99**	10.40	9.76	7.37
RMSE cold (cm)	5.49	**2.98**	6.75	5.99	4.23	4.24	6.32	6.56	***2.57***	5.73	3.74	3.17
% Bias overall	-71.76	**-3.73**	-32.12	3.40	-3.55	-16.14	6.05	6.29	***1.90***	3.63	1.86	0.01
% Bias warm	-87.54	**-1.74**	-23.44	-5.11	-2.23	-15.13	***-1.66***	-1.99	2.71	4.34	2.54	0.49
% Bias cold	**6.88**	-16.41	-90.76	-7.08	11.71	22.37	53.67	57.41	***-3.13***	-0.61	-2.20	-1.92
NSE overall	-0.70	***0.95***	0.81	0.82	0.89	0.89	0.89	0.89	**0.94**	0.88	0.91	0.93
NSE warm	-4.48	***0.88***	0.53	0.61	0.74	0.70	0.79	0.80	**0.82**	0.74	0.77	0.80
NSE cold	-1.36	**0.31**	-2.55	-0.93	0.04	-0.40	-2.12	-2.35	***0.48***	-0.79	0.25	0.21
RRMSE overall	0.82	**0.29**	0.72	0.49	0.39	0.34	0.11	0.11	***0.04***	0.58	0.39	0.39
RRMSE warm	0.83	***0.19***	0.31	0.32	0.28	0.22	0.23	**0.22**	0.35	0.37	0.34	0.37
RRMSE cold	0.80	**0.36**	0.96	0.62	0.47	0.39	1.36	1.41	***0.35***	0.73	0.43	0.42

Bold values indicate the best calibration values for each model, italics indicates the overall best calibration values for both models, black shading indicates best performing model for validation performance metrics, and gray shading indicates best performing model for reverse calibration and reverse validation performance metrics.

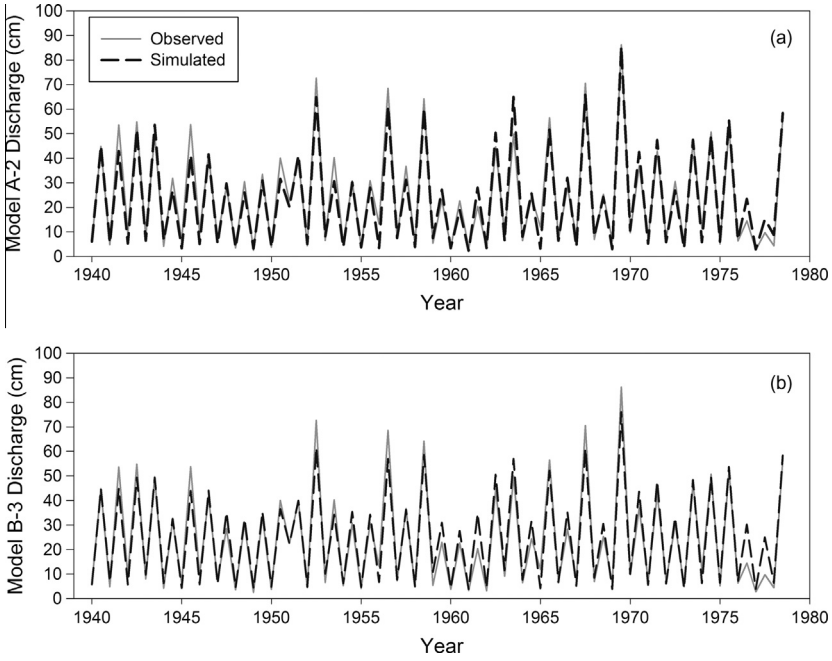

Fig. 4. Seasonal streamflow calibration results for (a) Model A-2, the modified Thornthwaite seasonal model with minimum average seasonal PRISM temperature inputs, and (b) Model B-3, the simple water balance model with WASMOD snow component and average seasonal PRISM temperature inputs. Calibration results for the same models, but using proxy inputs, are presented in detail in Devkota (2013).

Table 4
Performance metrics of the models using calibrated parameter values from shuffled complex evolution (SCE) method using the proxy reconstructed precipitation and temperatures for WY 1940–1980.

Performance metric (units)	Model A seasonal	Model B seasonal	Model A annual	Model B annual
R^2 overall	0.80	0.78	0.36	0.30
R^2 warm	0.41	0.34	--	--
R^2 cold	0.03	0.00	--	--
RMSE overall (cm)	10.71	10.07	15.17	15.86
RMSE warm (cm)	13.23	13.78	--	--
RMSE cold (cm)	6.46	3.57	--	--
% Bias overall	-19.33	0.77	0.95	16.63
% Bias warm	-12.61	1.06	--	--
% Bias cold	-68.23	-1.19	--	--
NSE overall	0.75	-0.29	0.33	0.11
NSE warm	0.32	0.31	--	--
NSE cold	-2.40	-0.04	--	--
RRMSE overall	0.65	0.56	0.39	0.59
RRMSE warm	0.37	0.63	--	--
RRMSE cold	0.85	0.49	--	--

RMSE = root mean squared error; NSE = Nash Sutcliffe Efficiency; RRMSE = relative root mean squared error. Black shading indicates best performing model for each set of calibration performance metrics.

The incremental parameter analysis for the Thornthwaite seasonal water balance model (Model A-2) showed that *rfactor*, *meltmax*, *Initial GS*, and *STC* are more sensitive than other parameters (Fig. 5a). For the simple water balance model with WASMOD and average seasonal temperatures (Model B-3), a_1, a, and c are more sensitive than other parameters (Fig. 5b).

We compared calibrated parameter values obtained using tree-ring precipitation and temperature with those obtained using PRISM data (Table 5). For Model A, values of *meltmax, rfactor, STC, GS*, and *Initial surface water* were similar, whereas values of *drofrac, Train,* and *Tsnow* were quite different. Model A was less sensitive to these latter parameters, with small changes in RMSE for different values of these parameters (Fig. 5a). For Model B, values of *a, b, c, d,* changed by only 5–13%, but the value of a_1 increased from 1.42 to 3.11°C when using tree-ring data and *Initial GS* changed from 1.70 cm to 13.18 cm (Table 5). RMSE did not change very much when a_1 and *Initial GS* were changed except when $a_1$was between about 0 and 0.4°C (Fig. 5b). Thus, these differences in values of a_1 and *Initial GS* were likely due to model insensitivity to these parameters rather than changes in the input variables.

Based on our analysis, the modified Thornthwaite model (Model A) was the best choice for reconstructing hydrologic parameters. The extended streamflow record from WY 1500 to 1980 (Fig. 6a) shows dry and wet episodes that were ranked (Table 6) based on duration, absolute magnitude, and absolute maximum (or peak) as suggested by Biondi et al. (2008). To demonstrate the utility of the mechanistic approach, the snow water equivalence (SWE; cm) remaining at the end of each WY and annual actual evapotranspiration (AET) were also plotted (Fig. 6b and c). Linear correlations between modeled SWE and measured SWE from five sites in the Upper West Walker River basin (Leavitt Meadows, Sonora Pass, Virginia Lakes, Virginia Lakes Ridge, and Willow Flat) between WY 1947 and WY 1980 had an average R^2 of 0.37 (range 0.30–0.51; Fig. 7).

Runoff anomalies, which were computed using the long-term mean of 31.6 cm, identified a total of 188 episodes, half positive (wet) and half negative (dry). The longest duration was the early 1900s pluvial, which in this reconstruction occurred during 1900–1923 (a 24-year wet interval), followed in terms of duration by a 12-year dry spell (1840–1851). Using the combination of duration, magnitude, and peak scores, the two strongest episodes were wet ones, i.e., the early 1900s pluvial and another 11-year positive spell from 1965 to 1975 (Table 6). The three strongest droughts, each with decadal-scale duration, occurred from the late 1700s to the mid-1800s. The 'Dust Bowl' period, which in this reconstruction corresponds to a negative interval from 1927 to 1933, was not among the strongest episodes, as it ranked in 21st position overall.

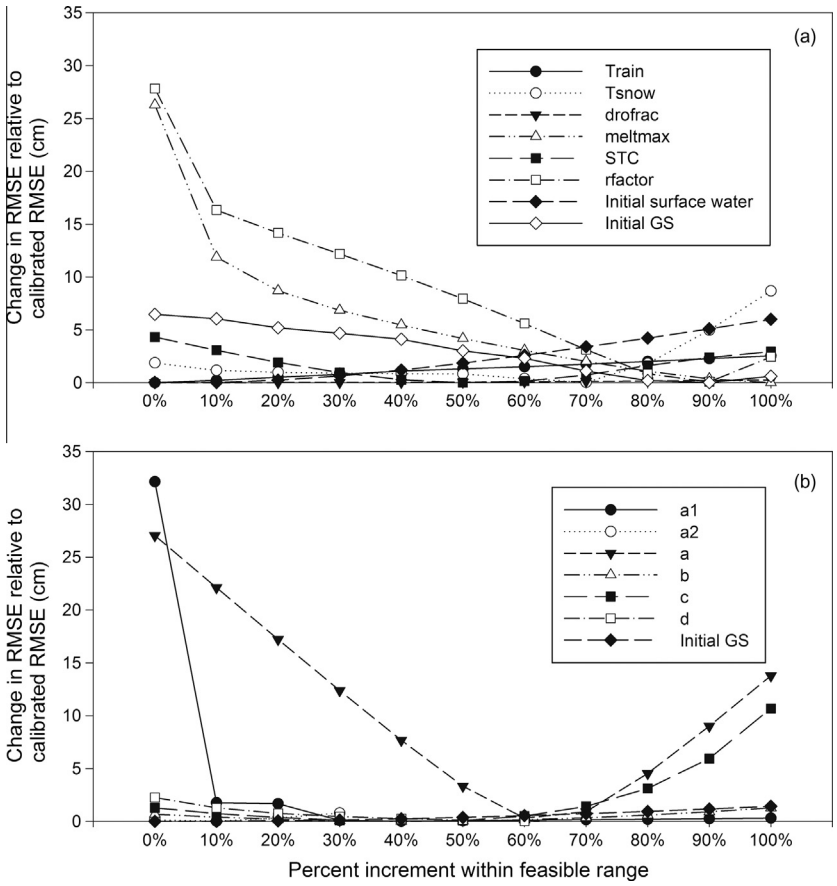

Fig. 5. Change in RMSE from SCE calibrated RMSE (cm) when parameter values are changed by 10 % over the feasible parameter space one at a time for (a) Model A-2 and (b) Model B-3.

Table 5
Parameter values obtained from Model A-2 and Model B-3 during the calibration period (WY 1940–1978) using PRISM precipitation and temperatures and from Model A and Model B during the calibration period (WY 1940–1980) using tree ring precipitation and temperatures.

Parameter (units)	Model A-2	Model A	Model B-3	Model B
T_{rain} (°C)	0.10	0.90		
T_{snow} (°C)	−7.09	−1.39		
drofrac	0.00	1.00		
meltmax	1.00	1.00		
STC (cm)	180.80	176.53		
rfactor	0.88	0.94		
a_1 (°C)			1.42	3.11
a_2 (°C)			0.29	0
a			0.64	0.59
b			0.41	0.43
c			0.37	0.30
d			0.63	0.70
Initial groundwater storage (*GS*) (cm)	184.41	190.44	1.70	13.18
Initial surface water (cm)	2.46	1.97		

4. Discussion

Comparison of our model results to previous studies indicate similar or better performance than those for Saito et al. (2008), Solander et al. (2010), and Gray and McCabe (2010), although additional factors such as geographic areas, calibration series, and tree-ring data could have affected differences in model performance. The R^2 of 0.80 for Model A, the modified Thornthwaite model with proxy precipitation and temperature inputs, was better than the correlation coefficient of 0.71 for the application of the Thornthwaite model by Gray and McCabe (2010). The calibration and validation R^2 values for the simple water balance model with WASMOD snow component and PRISM precipitation and temperatures (Table 3) were better than the R^2 obtained by Saito et al. (2008) and Solander et al. (2010), both also based on the simple water balance model using PRISM precipitation and temperature.

The calibrated value of parameter a (proportion of precipitation that becomes runoff) for Model B was about 0.6 (Table 5), meaning that about 60% of the incoming precipitation was treated as surface runoff. These values seem reasonable given the classification of the soils in the watershed as Soil Group D with slower infiltration rates and higher surface runoff (Natural Resources Conservation Service, 2006). For Model A, the estimation of runoff is more complex. The parameter drofrac for Model A represents the proportion of precipitation that becomes runoff for estimation of soil moisture storage and was calibrated as 0.00 for PRISM inputs and 1.00 for proxy inputs. The former value is not very realistic given the soils classification for the watershed. However, rfactor (the proportion of surplus available for runoff that becomes runoff) for both PRISM inputs and proxy inputs for Model A were large (0.88 and 0.94), which is more similar to a values for Model B.

The calibrated value of parameter *b* (proportion of infiltrated water that is evapotranspired) for Model B was 0.41 with PRISM inputs, and 0.48 with proxy inputs. Multiplication of parameter *b* and the proportion of water infiltrated $(1 - a)$ yielded estimated evapotranspiration of 26.2% and 25.4% of the total average annual precipitation (i.e., 19.2 cm and 18.6 cm) in Model B for PRISM inputs and proxy inputs, respectively, during the calibration period. These values are about half of the upper range of estimated

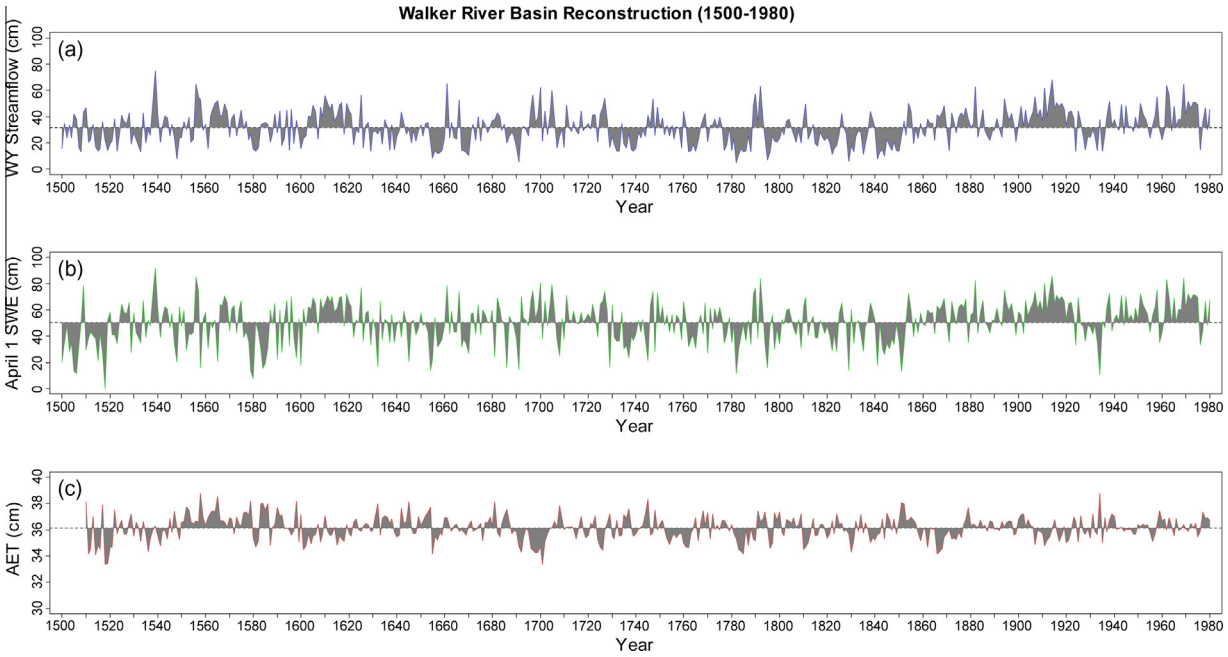

Fig. 6. Reconstructed hydrologic parameters for the upper West Walker River basin using Model A with proxy seasonal precipitation and temperature. Time series episodes with respect to the long-term average (horizontal dashed line) from 1500 to 1980 were highlighted with grayscale filling. (a) Water year streamflow (cm), (b) April 1 snow water equivalent (SWE, cm), (c) Actual evapotranspiration (AET, cm).

Table 6
List of the 10 strongest episodes for reconstructed water year streamflow (cm) from 1500 to 1980 (see Fig. 6a for a time series plot). The episode score was computed using anomalies (= deviations from the long-term mean) by ranking separately duration, magnitude, and peak (see text for details).

Start (year)	End (year)	Episode type	Duration (yrs)	Absolute magnitude (cm)	Absolute maximum (cm)	Score
1900	1923	Pos	24	314.0	36.12	563
1965	1975	Pos	11	162.3	33.12	555
1840	1851	Neg	12	172.6	23.88	546
1780	1788	Neg	9	149.1	26.58	543
1827	1837	Neg	11	110.1	25.48	541
1608	1617	Pos	10	147.5	24.12	539
1875	1882	Pos	8	82.9	30.82	535
1794	1802	Neg	9	92.8	24.78	533
1815	1825	Neg	11	119.3	20.38	529
1654	1660	Neg	7	115.2	23.38	527

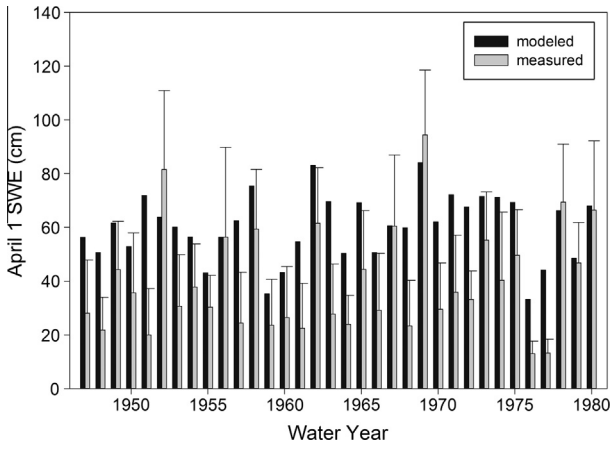

Fig. 7. Comparison of reconstructed April 1 snow water equivalent (SWE, cm) shown in Fig. 6b with average measured SWE for WY 1947–1980 derived from five sites in the upper West Walker River basin (Leavitt Meadows, Sonora Pass, Virginia Lakes, Virginia Lakes Ridge, and Willow Flat). Error bars represent the standard deviation of measured SWE.

mean annual actual evapotranspiration (31–40 cm) for Mono County, California which contains much of the upper West Walker River Basin (Sanford and Selnick, 2013).

A similar calibration parameter related to the estimation of evapotranspiration is not part of Model A; however, one of the advantages of the mechanistic approach is that we can examine model estimations of evapotranspiration to see if they are reasonable. Calibration results with proxy inputs show an average annual actual evapotranspiration (AET) estimate of 36.2 cm for Model A, whereas Model B had an average annual evapotranspiration estimate of 13.3 cm. For the entire reconstruction period (e.g., 1500–1980), average annual AET was 36.0 cm with Model A. Thus, the values determined with Model A were within the range estimated by Sanford and Selnick (2013). The process for estimating AET with Model A only relies on the proxy temperature inputs, whereas the estimation of evapotranspiration for Model B involves both the proxy temperature and proxy precipitation inputs.

Both the modified Thornthwaite model (Model A) and the simple water balance model with WASMOD snow component (Model B) performed well when comparing estimates of streamflow against gaged data in the upper West Walker River basin. Further,

both models have sensitive parameters that could be used to develop "what if" scenarios involving changes in streamflow due to climatic and non-climatic perturbations. For instance, the Thornthwaite model (Model A) could be used to test impacts of air temperature changes on estimated streamflow. However, Model A does not have a parameter that directly links to ET, so direct tests of ET impacts on streamflow from changes in vegetation or land cover cannot be quantified. The simple water balance model with WASMOD snow component (Model B) has model parameters that could be used to test changes in streamflow due to changes in watershed characteristics (Fig. 5b). However, the parameter that is directly related to ET (parameter b) is not very sensitive, though parameter a could be employed to test changes in streamflow due to changes in land cover and ET due to changes in runoff coefficient.

The performance statistics for the cold–wet season streamflows were consistently worse than the statistics for the warm–dry season streamflows. This is because cold–wet season precipitation accounts for most of the streamflow during the year, and most of the streamflow (an average of 86% of the water year total) occurs in the warm–dry season. Thus, the time lag between precipitation and streamflow likely results in the better model performance for the warm–dry season.

The high performance metrics, especially R^2, of the seasonal models may be partially due to the seasonal differences between warm–dry and cold–wet temperature and precipitation. To test this possibility, we estimated streamflows with the two models using annual proxy precipitation and temperature values. After SCE calibration with proxy inputs, the modified Thornthwaite model (Model A) had an R^2 of 0.36 with RMSE of 15.17 cm, and the simple water balance model with WASMOD inputs had an R^2 of 0.30 with RMSE of 15.86 cm. The R^2 values are similar to the statistics of the proxy precipitation regressions (Fig. 3a), so improved estimation of proxy inputs could improve model results. As another comparison, a linear regression model of water-year streamflow with the proxy inputs as predictors yielded significant coefficients with appropriate signs (positive for precipitation, negative for temperature) and an adjusted R^2 of 0.38.

Performance of our water balance approach is difficult to compare with other reconstructions because of different number of predictors, time periods, environmental factors, and ways to compute streamflow itself. Even though direct correlations between proxy records and their target climate variables were relatively low, the skill of the watershed model for reconstructing water-year flow was high. This suggests that our seasonal water-balance model is indeed combining the proxy inputs in a way that effectively enhances their ability to simulate observed streamflow. It should be noted that stochastic, rather than mechanistic, modeling of seasonal time series is complicated by the inherent periodicity in several statistical characteristics that invariably lead to stochastic models with periodic parameters (Salas and Obeysekera, 1992).

The advantage of a mechanistic approach, regardless of how well it fits error-prone observations, is the ability to assess sources of uncertainty to determine where model improvement is necessary, and to estimate hydrologic parameters not initially available (for yet another example, see Huang et al., 1996). In this application, it was possible to concurrently estimate streamflow, evapotranspiration, and April 1 snow water equivalent, which is used extensively to estimate water resources in the western US (Mote, 2003; Timilsena and Piechota, 2008). The sensitivity analysis and comparison of model estimations for streamflows, evaporatranspiration, and snow water equivalence were done to examine uncertainty in the parameter values and their physical interpretation. Since model over-fitting is a concern, we also performed a Monte Carlo analysis of model parameters, and the results, which are included in Devkota (2013), supported the sensitivity analysis shown in Fig. 5. Further research should explore model sensitivity and performance for estimating evapotranspiration, as well as model uncertainties related to input variability and parameter estimation.

Reconstructed streamflow episodes highlighted the pluvial period of the early twentieth century and the drought of the mid-1800s. The strongest wet and dry episodes from the streamflow reconstruction were similar to those for precipitation reconstructions in the same area (Biondi et al., 2008), but some similarity was also found with streamflow episodes reconstructed further east in Spring Valley, Nevada (Strachan et al., 2012), which lies in the central Great Basin near the Utah-Nevada border. In the Walker Basin reconstructed streamflow record, the dry period of the 'Dust Bowl' ranked far below the strongest episodes (21st place), similar to what was found for the nearby precipitation reconstruction (Biondi et al., 2008). The proxy record to the east (Strachan et al., 2012), which was for streamflow, indicated a much stronger 'Dust Bowl' drought, and additional evidence for spatial variability in the severity of drought episodes over the Great Basin has since been published (Biondi, 2014).

Additional comparisons with existing reconstructions included the Sacramento River streamflow developed by Meko et al. (2001), which was obtained from the NOAA/NGDC Paleoclimatology Program, and the summer (June to August) Palmer Drought Severity Index (PDSI; Alley, 1983) proxy series at intervals of 2.5° of latitude and longitude for all of North America (Cook and Krusic, 2004; Cook et al., 2004). The four PDSI grid nodes nearest to the Upper West Walker watershed are identified as 58 (117.5°W, 40.0°N), 59 (117.5°W, 37.5°N), 46 (120.0°W, 40.0°N), and 47 (120.0°W, 37.5°N). Linear correlations estimated using the Pearson's sample coefficient (r) for the period of overlap (1500–1977) showed that our streamflow reconstruction was better correlated ($r = 0.5$) with Great Basin proxy series (PDSI 58) than with Sierra Nevada ones ($r = 0.4$ with all other reconstructions).

5. Final remarks

The water balance approach we have outlined is suitable for semi-arid areas in the western United States and elsewhere when sparse instrumental data are available to parameterize common hydrologic models. In addition, estimated model parameters can be used to assess the overall hydrology of the watershed, and at the same time their values can be checked against existing information on water balance components. Given its limited input requirements, this approach performed adequately with proxy inputs (precipitation and temperature) to generate past streamflow, and could then be used to examine the influence of factors independent of streamflow (e.g., evapotranspiration, infiltration, and non-climatic factors such as land use changes) on its reconstructions. Improvements in proxy inputs, especially by means of localized air temperature reconstructions, should boost the performance of these mechanistic models and improve their insights into sensitivities and uncertainties.

Our water balance model could also be used with time series of future climatic variables generated from general circulation models (GCMs) to predict changes in water resources. These applications would provide concerned stakeholders with valuable information for water resources planning and management. Hence, we argue that the use of mechanistic modeling that only require proxy data inputs of precipitation and temperature for streamflow reconstructions has promise for extending streamflow records and improving estimates of watershed vulnerability to climate change, which could help water managers and policy makers to plan for mitigation of water resource deficits.

Acknowledgments

Research supported in part by the US National Science Foundation under Award No. AGS-P2C2-0823480 and under Cooperative Agreement EPS-0814372 to the Nevada System of Higher Education. F. Biondi was also supported in part by NSF Award No. BSC-1230329 and by a Charles Bullard Fellowship in Forest Research from Harvard Forest in Petersham, MA. L. Saito, F. Biondi, and R. Devkota were also partially supported by NSF Award No. EPS 0813472. We would like to thank Scotty Strachan for providing the PRISM temperature and precipitation data, as well as for his contribution, both in the field and in the laboratory, to the development of tree-ring proxy records. We also thank Alexes Garrett for her assistance with evaluating SWE estimates. The comments of two anonymous reviewers helped improve an earlier version of the manuscript.

Appendix A. Additional equations for models

A.1. Equations used in modified Thornthwaite model (Model A)

$$P_{snow,t} = P_{total,t} * \left[\frac{T_{rain} - T_t}{T_{rain} - T_{snow}}\right] 0.1 < T_{rain} < 5; \quad -10 < T_{snow} < 1 \tag{A.1}$$

$$P_{rain,t} = P_{total,t} - P_{snow,t} \tag{A.2}$$

$$DRO_t = P_{rain,t} \times drofrac; \quad 0.1 < drofrac < 1 \tag{A.3}$$

$$P_{remian,t} = P_{rain,t} - DRO_t \tag{A.4}$$

$$SMF_t = \left[\frac{T_t - T_{snow}}{T_{rain} - T_{snow}}\right] \times meltmax; \quad 0.0 < meltmax < 1 \tag{A.5}$$

$$SM_t = snoster * SMF_t \tag{A.6}$$

$$PET_t = 13.97 \times d \times D^2 \times W_t \tag{A.7}$$

$$W_t = \frac{4.95 * e^{0.62*T}}{100} \tag{A.8}$$

$$P_{liquid,t} = P_{remain,t} + SM_t \tag{A.9}$$

$$ST_t = ST_{t-1} - \left[abs(P_{liquid,t} - PET_t) \times \left\{\frac{ST_{t-1}}{STC}\right\}\right] \tag{A.10}$$

$$STW_t = ST_{t-1} - ST_t \tag{A.11}$$

$$AET_t = \left\{\begin{array}{l} PET_t \text{ for } P_{liquid,t} \geqslant PET_t \\ P_{liquid,t} + STW_t \text{ for } P_{liquid,t} < PET_t \end{array}\right\} \tag{A.12}$$

$$S_t = ST_t - STC \tag{A.13}$$

$$RO_t = (S_t + Initial\ surface\ water_t) * rfactor + DRO_t \tag{A.14}$$

$$Intial\ surface\ water_t = S_{t-1} + Intial\ surface\ water_{t-1} - RO_{t-1} \tag{A.15}$$

where $P_{total,t}$ = total precipitation in season t (mm); $P_{rain,t}$ = total precipitation in a season that is rainfall (mm); $P_{snow,t}$ = total snowfall in a season(mm); T_{snow} = temperature threshold below which all seasonal precipitation is considered as snow (°C); T_{rain} = temperature threshold above which all seasonal precipitation is considered as rainfall (liquid precipitation) (°C); RO_t = runoff (mm); DRO_t = direct runoff in a season (mm); drofrac = direct runoff coefficient; $P_{remain,t}$ = total remaining precipitation (mm); SMF_t = snowmelt fraction; meltmax = maximum melt rate; SM_t = total snowmelt in a season (mm); PET_t = potential evapotranspiration using Hamon (1961) equation (mm); d = number of days in a season; D = mean seasonal hours of daylight in units of 12 h; W_t = saturated water vapor density (gm/m^3); ST_t = Soil moisture storage in season; STW_t = soil-moisture storage withdrawal; STC = soil moisture storage capacity (mm); S_t = surplus; $rfactor$ = runoff factor.

A.2. Equations used in simple water balance model with WASMOD snow component (Model B)

$$SR_t = a * P_t; 0 \leqslant a \leqslant 0.9 \tag{A.16}$$

$$ET_t = b * I_t; 0 \leqslant b \leqslant 1 \tag{A.17}$$

$$BF_t = c * GS_{t-1}; 0 \leqslant c \leqslant 1 \tag{A.18}$$

$$GF_t = d * GS_{t-1}; 0 \leqslant d \leqslant 1 \tag{A.19}$$

$$I_t = P_t - SR_t \tag{A.20}$$

$$DP_t = I_t - ET_t \tag{A.21}$$

$$GS_t = GS_{t-1} + DP_t - BF_t - GF_t \tag{A.22}$$

$$Q_t = SR_t + BF_t \tag{A.23}$$

where a = fraction of snowmelt and rain that becomes runoff; b = fraction of infiltration that becomes evapotranspiration; c = fraction of groundwater storage that becomes base flow; d = fraction of groundwater storage that becomes groundwater flow; Sn_t = seasonal snowfall component (cm); P_t = seasonal precipitation (cm); a_1 = temperature threshold above which snowfall starts to become rainfall; a_2 = temperature threshold above which snowmelt begins; R_t = seasonal rainfall (cm); M_t = seasonal snowmelt (cm); SR_t = surface runoff for season t (cm); ET_t = evapotranspiration for season t (cm); BF_t = base flow for season t (cm); GF_t = groundwater flow for season t (cm); I_t = infiltration for season t; DP_t = deep percolation for season t; GS_t = groundwater storage for season t v; GS_{t-1} = groundwater storage for previous season (or initial boundary condition for season 1) (cm); P_t = precipitation for season t (cm); and Q_t = streamflow for season t (cm).

References

Alley, W.M., 1983. The palmer drought severity index: limitations and assumptions. J. Climate Appl. Meteorol. 23, 1100–1109.

Alley, W.M., 1984. On the treatment of evapotranspiration, soil moisture accounting, and aquifer recharge in monthly water balance models. Water Resour. Res. 20 (8), 1137–1149.

Anderson, M.T., Woosley Jr. L.H., 2005. Water availability for the Western United States—key scientific challenges., U.S. Geological Survey Circular 1261, 85 p. <http://pubs.usgs.gov/circ/2005/circ1261/pdf/C1261.pdf>.

Bardossy, A., Singh, S.K., 2008. Robust estimation of hydrological model parameters. Hydrol. Earth Syst. Sci. 12, 1273–1283.

Biondi, F., 2014. Dendroclimatic reconstruction at km-scale grid points: a case study from the Great Basin of North America. J. Hydrometeorol. 15 (2), 891–906.

Biondi, F., Kozubowski, T.J., Panorska, A.K., Saito, L., 2008. A new stochastic model of episode peak and duration for eco-hydro-climatic applications. Ecol. Model. 211, 383–395.

Bradley, R.S., 1999. Paleoclimatology, second ed. Academic Press, San Diego, CA.

Brath, A., Montanari, A., Toth, E., 2004. Analysis of the effects of different scenarios of historical data availability on the calibration of a spatially-distributed hydrological model. J. Hydrol. 219, 232–253.

Cayan, D.R., Das, T., Pierce, D.W., Barnett, T.P., Tyree, M., Gershunov, A., 2010. Future dryness in the Southwest US and the hydrology of the early 21st century drought. Proc. Natl. Acad. Sci. U.S.A. 107 (50), 21271–21276.

Cook, E.R., Kairiukstis, L.A., 1990. Methods of Dendrochronology: Applications in the Environmental Sciences. Kluwer Academic Publishers, Dordrecht, the Netherlands.
Cook, E.R., Krusic, P.J., 2004. The North American Drought Atlas. Lamont-Doherty Earth Observatory, Palisades, New York, USA. <http://iridl.ldeo.columbia.edu/SOURCES/.LDEO/.TRL/.NADA2004/.pdsi-atlas.html>.
Cook, E.R., Woodhouse, C.A., Eakin, C.M., Meko, D.M., Stahle, D.W., 2004. Long-term aridity changes in the western United States. Science 306 (5), 1015–1018.
Daly, C., Neilson, R.P., Phillips, D.L., 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. J. Appl. Meteorol. 33 (2), 140–158.
Devkota, R., 2013. Mechanistic water balance approach for reconstructing past streamflow using tree ring (proxy) records. Master's thesis. Reno: University of Nevada Reno. 99 p.
Duan, Q.Y., Gupta, V.K., Sorooshian, S., 1992. Effective and efficient global optimization for conceptual rainfall–runoff models. Water Resour. Res. 28 (4), 1015–1031.
Duan, Q.Y., Sorooshian, S., Gupta, V.K., 1994. Optimal use of the SCE global optimization method for calibrating watershed models. J. Hydrol. 158 (3–4), 265–284.
Eckhardt, K., Fohrer, N., Frede, H.G., 2005. Automatic model calibration. Hydrol. Process. 19 (3), 651–658.
Fiering, M.B., 1967. Streamflow Synthesis. Harvard University Press, Cambridge.
Frank, D., Wilson, R., Esper, J., 2005. Synchronous variability changes in alpine temperature and tree-ring data over the past two centuries. Boreas 34 (4), 498–505.
Fritts, H.C., 1976. Tree Rings and Climate. Academic Press, London, United Kingdom, ISBN: 13: 978-1930665392.
Gray, S.T., McCabe, G.J., 2010. A combined water balance and tree ring approach to understanding the potential hydrologic effects of climate change in the central Rocky Mountain region. Water Resour. Res., w05513. http://dx.doi.org/10.1029/2008wr007650.
Gupta, H.V., Sorooshian, S., Yapo, P.O., 1999. Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. J. Hydrol. Eng. 4 (2), 135–143. http://dx.doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135).
Hamon, W.R., 1961. Estimating potential evapotranspiration. J. Hydraul. Div. Proc. Am. Soc. Civil Eng. 83 (HY 3), 107–120.
Hanson, C.L., 2001. Long-term precipitation database, Reynolds Creek Experimental Watershed, Idaho, United States. Water Resour. Res. 37 (11), 2831–2834.
Hill, M., 1975. Geology of the Sierra Nevada. University of California Press, Berkeley.
Horton, G., 1996. Walker River Chronology: A Chronological History of the Walker River and Related Water Issues. Nevada Division of Water Planning, Carson City.
Huang, J., van den Dool, H.M., Georgakakos, K.P., 1996. Analysis of model-calculated soil moisture over the United States (1931–1993) and applications to long-range temperature forecasts. J. Clim. 9, 1350–1362.
Kattelmann, R., 2012. East Walker River Watershed Assessment. Eastern Sierra Land Trust, Bishop, CA. 177 p. <http://inyo-monowater.org/wp-content/uploads/2011/09/E-Walker-Assessment-FINAL.pdf> (May 13, 2013).
Klemes, V., 1986. Operational testing of hydrological simulation models. Hydrol. Sci. J. 31, 13–24.
McCabe, G.J., Markstrom, S.L., 2007. A monthly water-balance model driven by a graphical user interface. Open-file report 2007–1088. U.S. Geological Survey, Reston, VA.
McCabe, G.J., Wolock, D.M., 1999. General-circulation-model simulations of future snowpack in the Western United States. J. Am. Water Resour. Assoc. 35 (6), 1473–1484.
Meko, D., Graybill, D.A., 1995. Tree-ring reconstruction of upper Gila River discharge. Water Resour. Bull. 31 (4), 605–616.
Meko, D.M., Therrell, M.D., Baisan, C.H., Hughes, M.K., 2001. Sacramento River flow reconstructed to A.D. 869 from tree rings. J. Am. Water Resour. Assoc. 37 (4), 1029–1039.
Milly, P.C.D., Dunne, K.A., Vecchia, A.V., 2005. Global pattern of trends in streamflow and water availability in a changing climate. Nature 438 (7066), 347–350.
Mote, P.W., 2003. Trends in snow water equivalent in the Pacific Northwest and their climatic causes. Geophys. Res. Lett. 30 (12), 1601–1604.
Natural Resources Conservation Service (NRCS), 2009. National Engineering Handbook. USDA-NRCS, Engineering Division. U.S. Gov. Print. Office, Washington, DC. Part 630, Section 4, Chapter 7.
Nelder, J.A., Mead, R., 1965. A simplex method for function minimization. Comput. J. 7 (4), 308–313.
Pavelsky, T.M., Kapnick, S., Hall, A., 2011. Accumulation and melt dynamics of snowpack from a multiresolution regional climate model in the Central Sierra Nevada, California. J. Geophys. Res. 116, D16115. http://dx.doi.org/10.1029/2010JD015479.
Rango, A., Martinec, J., 1995. Revisiting the degree-day method for snowmelt computations. Water Resour. Bull. 31, 657–669.
Reed, R.D., 1933. Geology of California. American Association of Petroleum Geologists, Tulsa, OK.
Remick, C.C., 2003. How local governments respond to a proposed purchase of water rights by a federal agency: a case study in the Walker River Basin. M.S Thesis, University of Nevada, Reno, Nevada.
Saito, L., Biondi, F., Salas, J.D., Panorska, A.K., Kozubowski, T.J., 2008. A watershed modeling approach to streamflow reconstruction from tree-ring records. Environ. Res. Lett. 3 (2), 024006. http://dx.doi.org/10.1088/1748-9326/3/2/024006.
Salas, J.D., Obeysekera, J., 1992. Conceptual basis of seasonal streamflow time series models. ASCE J. Hydraul. Eng. 118 (8), 1186–1194.
Sanford, W.E., Selnick, D.L., 2013. Estimation of evapotranspiration across the conterminous United States using a regression with climate and land-cover data. J. Am. Water Resour. Assoc. 49 (1), 217–230.
Seager, R., Ting, M., Held, I., Kushnir, Y., Lu, J., Vecchi, G., Huang, H.P., Harnik, N., Leetmaa, A., Lau, N.C., Li, C., Velez, J., Naik, N., 2007. Model projections of an imminent transition to a more arid climate in Southwestern North America. Science 316 (5828), 1181–1184.
Serreze, M.C., Clark, M.P., Frei, A., 2001. Characteristics of large snowfall events in the montane Western United States as examined using Snowpack Telemetry (SNOTEL) data. Water Resour. Res. 37 (3), 675–688.
Sieber, A., Uhlenbrook, S., 2005. Sensitivity analyses of a distributed catchment model to verify the model structure. J. Hydrol. 310 (1–4), 216–235.
Singh, V.P., Woolhiser, D.A., 2002. Mathematical modelling of watershed hydrology. J. Hydrol. Eng. 7 (4), 270–292.
Solander, K., Saito, L., Biondi, F., 2010. Streamflow simulation using a water-balance model with annually-resolved inputs. J. Hydrol. 387 (1–2), 46–53.
Sorooshian, S., Duan, Q., Gupta, V.K., 1993. Calibration of conceptual rainfall-runoff models using global optimization: application to the Sacramento soil moisture accounting model. Water Resour. Res. 29 (4), 1185–1194.
Stewart, I.T., 2009. Changes in snowpack and snowmelt runoff for key mountain regions. Hydrol. Process. 23 (1), 78–94.
Stewart, I.T., Cayan, D.R., Dettinger, M.D., 2004. Changes in snowmelt runoff timing in Western North America under a 'business as usual' climate change scenario. Climatic Change 62 (1–3), 217–232.
Strachan, S., Biondi, F., Leising, J., 2012. 550-year reconstruction of streamflow variability in Spring Valley, Nevada. J. Water Resour. Plann. Manage. 138 (4), 326–333.
Thornthwaite, C.W., 1948. Approach toward a rational classification of climate. Geogr. Rev. 38 (1), 55–94.
Thornthwaite, C.W., Mather, J.R., 1955. The Water Balance. Drexel Institute of Technology Laboratory of Climatology, N J Publications in Climatology, Centerton, NJ, p. 104.
Thyer, M., Kuczera, G., Bates, B.C., 1999. Probabilistic optimization for conceptual rainfall-runoff models: a comparison of the shuffled complex evolution and simulated annealing algorithms. Water Resour. Res. 35 (3), 767–773.
Timilsena, J., Piechota, T.C., 2008. Regionalization and reconstruction of snow water equivalent in the upper Colorado River basin. J. Hydrol. 352, 94–106.
Tingley, M.P., Huybers, P., 2010. A Bayesian algorithm for reconstructing climate anomalies in space and time. Part II: comparison with the regularized expectation–maximization algorithm. J. Clim. 23, 2782–2800.
Vanhaute, W.J., Vandenberghe, S., Scheerlinck, K., De Beats, B., Verhoest, N.E.C., 2012. Calibration of the modified Bartlett-Lewis model using global optimization techniques and alternative objective functions. Hydrol. Earth Syst. Sci. 16 (3), 873–891. http://dx.doi.org/10.5149/hess-16-873-2012.
Vittori, J.C., 2011. Developing a watershed modeling approach for reconstructing past streamflow in the upper walker river basin, California. M.S. Thesis, University of Nevada, Reno, Reno, Nevada.
Vogel, R.M., Sankarasubramanian, A., 2003. Validation of a watershed model without calibration. Water Resour. Res. 39 (10), 1292. http://dx.doi.org/10.1029/2002WR001940.
Wahl, E.R., Smerdon, J.E., 2012. Comparative performance of paleoclimate field and index reconstructions derived from climate proxies and noise-only predictors. Geophys. Res. Lett. 39 (6), L06703. http://dx.doi.org/10.1029/2012/GL051086.
Wilhite, D.A., 2000. Drought: A Global Assessment: Routledge Hazards and Disasters Series. Routledge; London, New York.
Williams, K.S., Tarboton, D.G., 1999. The ABC's of snowmelt: a topographically factorized energy component snowmelt model. Hydrol. Process. 13 (12–13), 1905–1920.
Wolock, D.M., McCabe, G.J., 1999. Explaining spatial variability in mean annual runoff in the conterminous United States. Climate Res. 11, 149–159.
Woodhouse, C.A., Lukas, J.J., 2006. Multi-century tree-ring reconstructions of Colorado streamflow for water resource planning. Climatic Change 78 (2–4), 293–315. http://dx.doi.org/10.1111/j.1752-1688.1999.tb04219.x.
Wurbs, R.A., 1998. Dissemination of generalized water resources models in the United States. Water Int. 23 (3), 190–198.
Xiao, Z., Liang, S.L., Wang, J.D., Song, J.L., Wu, X.Y., 2009. A temporally-integrated inversion method for estimating leaf area index from MODIS data. IEEE Trans. Geosci. Remote Sens. 47 (8), 2536–2545. http://dx.doi.org/10.1109/TGRS.2009.2015656.
Xu, C.Y., 2002. WASMOD – The Water and Snow Balance Modeling System. In: Singh, Vijay P., Frevert, Donald (Eds.), Mathematical Models of Small Watershed Hydrology and Applications. Water Resources Publication LLC., Highlands Ranch, CO, USA, pp. 555–590 (Chapter 17).